Álgebra Linear I

Sonia Elena Palomino Castro Bean
Daniel Noberto Kozakevich

2ª Edição
Florianópolis, 2011

Governo Federal

Presidente da República: Dilma Vana Rousseff

Ministro de Educação: Fernando Haddad

Coordenador Nacional da Universidade Aberta do Brasil: Celso Costa

Universidade Federal de Santa Catarina

Reitor: Alvaro Toubes Prata

Vice-Reitor: Carlos Alberto Justo da Silva

Secretário de Educação a Distância: Cícero Barbosa

Pró-Reitora de Ensino de Graduação: Yara Maria Rauh Müller

Pró-Reitora de Pesquisa e Extensão: Débora Peres Menezes

Pró-Reitor de Pós-Graduação: Maria Lúcia de Barros Camargo

Pró-Reitor de Desenvolvimento Humano e Social: Luiz Henrique Vieira Silva

Pró-Reitor de Infra-Estrutura: João Batista Furtuoso

Pró-Reitor de Assuntos Estudantis: Cláudio José Amante

Centro de Ciências da Educação: Wilson Schmidt

Centro de Ciências Físicas e Matemáticas: Tarciso Antônio Grandi

Centro de Filosofia e Ciências Humanas: Roselane Neckel

Curso de Licenciatura em Matemática na Modalidade à Distância

Coordenação de Curso: Neri Terezinha Both Carvalho

Coordenação de Tutoria: Jane Crippa

Coordenação Pedagógica/CED: Roseli Zen Cerny

Coordenação de Ambientes Virtuais/CFM: Nereu Estanislau Burin

Comissão Editorial

Antônio Carlos Gardel Leitão

Albertina Zatelli

Elisa Zunko Toma

Igor Mozolevski

Luiz Augusto Saeger

Roberto Corrêa da Silva

Ruy Coimbra Charão

Laboratório de Novas Tecnologias - LANTEC/CED

Coordenação Pedagógica

Coordenação Geral: Andrea Lapa, Roseli Zen Cerny

Núcleo de Formação: Nilza Godoy Gomes

Núcleo de Pesquisa e Avaliação: Claudia Regina Flores

Núcleo de Criação e Desenvolvimento de Materiais

Design Gráfico

Coordenação: Laura Martins Rodrigues, Thiago Rocha Oliveira

Projeto Gráfico Original: Diogo Henrique Ropelato, Marta Cristina Goulart Braga, Natal Anacleto Chicca Junior

Redesenho do Projeto Gráfico: Laura Martins Rodrigues, Thiago Rocha Oliveira

Diagramação: Kallani Maciel Bonelli, Karina Silveira

Ilustrações: Gabriela Dal Toé Fortuna, Kallani Maciel Bonelli

Capa: Rafael Naravan Kienen

Design Instrucional

Coordenação: Elizandro Maurício Brick

Design Instrucional: Maria Carolina Machado Magnus

Revisão Gramatical: Daniela Piantola, Raquel Coelho, Tony Roberson De M. Rodrigues

Copyright © 2011, Universidade Federal de Santa Catarina/CFM/CED/UFSC
Nenhuma parte deste material poderá ser reproduzida, transmitida e gravada, por qualquer meio eletrônico, por fotocópia e outros, sem a prévia autorização, por escrito, da Coordenação Acadêmica do Curso de Licenciatura em Matemática na Modalidade à Distância.

Ficha Catalográfica

K88a Kozakevich, Daniel
Álgebra Linear I / Daniel Norberto Kozakevich, Sonia Elena Palomino Castro Bean. – 2. ed. – Florianópolis : UFSC/EAD/CED/CFM, 2011.
250 p. : il. ; grafs. , tabs.

Inclui bibliografia
UFSC. Licenciatura em Matemática na Modalidade a Distância
ISBN 978-85-8030-023-9

1. Álgebra linear. I. Bean, Sonia Elena P. Castro. II. Título.
CDU 512.64

Catalogação na fonte pela Biblioteca Universitária da UFSC

Sumário

Apresentação

A Álgebra Linear é o estudo dos espaços vetoriais e das transformações lineares definidas entre eles. Quando os espaços têm dimensões finitas, as transformações lineares podem ser representadas por matrizes. Também com matrizes podem ser representadas as formas bilineares e, mais particularmente, as formas quadráticas. Assim a Álgebra Linear, além de vetores e transformações lineares, lida também com matrizes e formas quadráticas. São numerosas e bastante variadas as situações, em Matemática e em suas aplicações, onde esses objetos se apresentam. Daí a importância central da Álgebra Linear no ensino da Matemática.

Neste livro se introduzem os conceitos da Álgebra Linear, desde os mais simples, que são as matrizes, até os mais abstratos, quando se trata do estudo de espaços vetoriais. Todos esses conceitos são apresentados, dentro do possível, de uma forma acessível, ajudando a compreensão com muitos exemplos, exercícios resolvidos e propostos. Também, com o objetivo de facilitar a compreensão do conteúdo, colocamos alguns tópicos com detalhes e justificações que usualmente não são expostos nos livros tradicionais.

Este texto pretende fornecer conceitos suficientes para que os estudantes consigam ter acesso ao nível dos livros avançados. Isto não significa deixar para trás as possibilidades que oferece a utilização de um software matemático ou ignorar as aplicações, no favor de uma exclusiva e única compreensão da Matemática. Significa que se pretende, principalmente, que o leitor obtenha uma compreensão global dos conceitos (como por exemplo, que a multiplicação de uma matriz por um vetor pode ser entendida como a aplicação de uma transformação linear) e também consiga acompanhar as provas e demonstrações.

O primeiro capítulo trata de Matrizes e Aplicações. No segundo capítulo, se estudam os Sistemas Lineares, começando com uma breve revisão dos conceitos da Geometria Analítica, para poder entender em uma forma geométrica como é que tais sistemas podem ser caracterizados. No terceiro capítulo define-se Espaço Vetorial, um conceito básico da Álgebra Linear que proporciona unidade e precisão aos assuntos essenciais da Matemática. E finalmente,

o quarto capítulo introduz a noção de Transformação Linear e as relações que existem entre transformações lineares e matrizes.

Embora a apresentação esteja focalizada sobre os principais tópicos da Álgebra Linear, não pressupõe que os estudantes possuam desde o início uma prática em trabalhar com conceitos que demandem certos níveis de abstração, ainda que desejável. Em lugar disso, esta atividade é estimulada através dos muitos exemplos e exercícios que diferem das verificações rotineiras ou uso de técnicas de resolução. O objetivo está colocado principalmente em desenvolver, sendo o material usual de um curso de graduação, o nível de maturidade matemática de um estudante da Licenciatura de Matemática.

Sonia Elena Palomino Castro Bean

Daniel Noberto Kozakevich

Capítulo 1

Matrizes

Capítulo 1

Matrizes

Ao finalizar o estudo deste Capítulo você será capaz de identificar alguns tipos de matrizes, fazer operações e provar propriedades e teoremas sobre matrizes. Também, será capaz de compreender e aplicar o conceito de matrizes em situações reais.

1.1 Matriz

As matrizes são estruturas matemáticas que podem ser encontradas em muitos problemas do nosso dia-a-dia. Por isso, neste capítulo, iniciaremos o estudo das matrizes com um problema vindo do nosso cotidiano.

Problema 1. Já pensou que a temperatura que temos em cada estação do ano pode ser registrada dia a dia e hora a hora (e até minuto a minuto!), com ajuda de dispositivos especiais? Isso é feito pelo Instituto de Meteorologia de cada uma das regiões. Considere a seguinte situação:

As temperaturas de cinco cidades brasileiras nas primeiras horas da manhã de um determinado dia (durante o inverno) foram registradas da forma seguinte:

- **Cidade n° 1**: São Joaquim (SC) às 3 horas da manhã apresenta -3 graus centígrados;
- **Cidade n° 2**: Rio de Janeiro (RJ) às 5 horas da manhã apresenta 14 graus centígrados;
- **Cidade n° 3**: Turvo (SC) às 7 horas da manhã apresenta 5 graus centígrados;
- **Cidade n° 4**: Florianópolis (SC) às 9 horas da manhã apresenta 16 graus centígrados;

- **Cidade n° 5**: São Luis (MA) às 11 horas da manhã apresenta 20 graus centígrados.

Essas informações podem ser arranjadas em tabelas de várias formas, como as que apresentamos a seguir:

Cidade	Temperatura (°C)
1	-3
2	14
3	5
4	16
5	20

Cidade	Hora
1	3
2	5
3	7
4	9
5	11

Hora	Temperatura (°C)
3	-3
5	14
7	5
9	16
11	20

Hora	Cidade
3	1
5	2
7	3
9	4
11	5

Observe que dessa forma as informações estão dispostas em forma vertical, mas também podemos colocar as mesmas informações em forma horizontal.

Pergunta 1. De que forma podem ser arranjados os dados acima de modo a estarem dispostos horizontalmente?

Se considerarmos H como sendo a hora e T a temperatura da cidade, então, a terceira tabela pode ser disposta da seguinte maneira:

H	3	5	7	9	11
T (°C)	-3	14	5	16	20

Deixamos de atividade para você completar essa disposição horizontal no caso das outras tabelas.

Continuando com o **Problema 1**, suponhamos que por algum motivo é do nosso interesse os dados do arranjo dado pela última tabela.

Assim, podemos formular o seguinte:

Em cinco cidades brasileiras, em determinadas horas, foram registradas as seguintes temperaturas:

H	T (°C)
3	-3
5	14
7	5
9	16
11	20

Observação. A mesma informação poderia ter sido colocada da seguinte forma:

H	3	5	7	9	11
T (°C)	-3	14	5	16	20

Os dois jeitos de arranjar os dados estão nos fornecendo o que denominaremos como **Matriz**.

1.1.1 Definição de matriz

É de nosso interesse trabalhar apenas com números reais neste Livro, assim sendo tudo o que será definido mais adiante, no caso das matrizes ou vetores, será com elementos reais (mais adiante você terá a possibilidade de trabalhar com números complexos também!).

Uma matriz é um arranjo de números, símbolos, letras, etc., dispostos em linhas e colunas.

1.1.2 Ordem de uma matriz

As matrizes geralmente são denotadas por letras maiúsculas e seus elementos por letras minúsculas. Se uma matriz possui m linhas e n colunas diremos que a matriz tem ordem $m \times n$.

Exemplo 1. Denominemos por A e B as duas matrizes definidas no **Problema 1** e na **Pergunta 1**, respectivamente. Assim:

$$A = \begin{bmatrix} 3 & -3 \\ 5 & 14 \\ 7 & 5 \\ 9 & 16 \\ 11 & 20 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 3 & 5 & 7 & 9 & 11 \\ -3 & 14 & 5 & 16 & 20 \end{bmatrix}.$$

A matriz A tem 5 linhas e 2 colunas, ou seja, é de ordem 5×2; já a matriz B tem 2 linhas e 5 colunas e é de ordem 2×5.

O elemento da 2ª linha e 2ª coluna da matriz A é igual a 14, ou seja:

$$a_{22} = 14.$$

O elemento da 1ª linha e 4ª coluna da matriz B é igual a 9, isto é:

$$b_{14} = 9.$$

Quando uma matriz é obtida por algum problema específico (como o explicitado no **Problema 1**), é possível fornecer alguma interpretação aos seus elementos.

Por exemplo, as matrizes A e B do **Exemplo 1** com elementos $a_{22} = 14$ e $b_{14} = 9$ podem ser interpretadas da seguinte forma:

*"No **segundo** horário (5 horas da manhã) o **segundo** valor da temperatura (no Rio de Janeiro) é 14 graus".*

"São 9 horas da manhã quando a temperatura em Florianópolis é 16 graus".

E, claro, após fornecermos todas as interpretações podemos fazer algumas conclusões:

- Eu gosto do frio, portanto irei para São Joaquim no inverno.
- Não, não gosto de tanto frio, por isso no inverno ficarei no Rio de Janeiro.

Bom, você deve estar se perguntando: onde está a matemática nesse papo todo? Se estiver fazendo esse tipo de questionamento está indo por um bom caminho, pois a matemática, por incrível que pareça, está presente em muitas situações! E é isso que esperamos mostrar ao longo deste material!

Lembrete. A partir de agora, serão apresentados vários exercícios que pediremos para você resolver.

Agora verifique se você está acompanhando as discussões que fizemos, resolvendo os seguintes exercícios.

Exercício 1. Coloque mais alguma condição no Problema 1 para construir uma matriz de ordem 3 x 5.

Dica: Imagine que os dados são colhidos durante 3 dias.

Exercício 2. Será que você pode imaginar e criar um problema do seu cotidiano diferente do dado acima para chegar a uma matriz?

Dada uma linha i e uma coluna j de uma matriz A, o elemento na posição (i, j) será denotado por a_{ij}. Assim, uma matriz com $m \times n$ elementos pode ser escrita na seguinte forma estendida:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

Também podemos colocá-la na forma abreviada:

$$A = \left[a_{ij} \right]_{m \times n}$$

Assim, a matriz A de ordem $m \times n$ possui $m \cdot n$ elementos da forma a_{ij} com $i = 1, \ldots, m$ e $j = 1, \ldots, n$.

Alguns livros denotam a matriz A de elementos a_{ij} na forma

$$A = (a_{ij})_{m \times n}.$$

Muitas vezes é fornecida uma lei de formação para obtermos os elementos de uma matriz. Por exemplo, se $A = \left[a_{ij} \right]_{2 \times 3}$ com $a_{ij} = i + j$, com $m = 2$ e $n = 3$, estaremos construindo a seguinte matriz A:

$$A = \begin{bmatrix} 1+1 & 1+2 & 1+3 \\ 2+1 & 2+2 & 2+3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}.$$

Exemplo 2. Vamos obter a matriz $B=(b_{ij})_{3\times 4}$, de ordem 3×4, cujos elementos são da forma

$$b_{ij}=\begin{cases} i^j, & i=1,2 \\ 0, & i=3 \end{cases}.$$

Solução. Observe que não há nenhuma condição para os índices j, isto é, j está variando conforme o número de colunas que a matriz tem. Já na 3ª linha $(i=3)$ todos os elementos serão nulos. Assim sendo, a matriz B é dada por:

$$B=\begin{bmatrix} 1^1 & 1^2 & 1^3 & 1^4 \\ 2^1 & 2^2 & 2^3 & 2^4 \\ 0 & 0 & 0 & 0 \end{bmatrix}=\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 8 & 16 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

1.2 Tipos de Matrizes

1.2.1 Matriz Retangular

São denominadas assim aquelas matrizes cujo número de linhas é diferente do número de colunas. Por exemplo:

$$A=\begin{bmatrix} 1 & -1 \\ 0 & 9 \\ 5 & -2 \end{bmatrix}\ B=\begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 7 & -2 & 3 & 8 & 0 \\ -1 & 3 & 3 & 2 & 6 \\ 3 & 5 & 0 & 0 & 9 \end{bmatrix}\ \text{e}\ C=\begin{bmatrix} 0 & 0 & 1 \\ -1 & 3 & 9 \end{bmatrix},$$

e podem ser colocadas na forma $A=[\ ]_{3\times 2}$, $B=[\ ]_{4\times 5}$ e $C=[\ ]_{2\times 3}$. No que segue podemos omitir a ordem na representação da matriz toda vez que ela venha dada na forma estendida.

1.2.2 Matriz Linha

A matriz linha é uma matriz que tem apenas uma linha. Por exemplo:

$$L=[1\quad 2\quad 3\quad 4] \qquad\qquad M=(0\quad 0\quad 1\quad 8).$$

Observação. É comum colocarmos vetores no plano e no espaço como matrizes linha entre parênteses, onde os elementos estão separados por vírgula. Exemplo: $(0,0,1,8)$.

1.2.3 Matriz Coluna

A matriz coluna é uma matriz que tem apenas uma coluna. Por exemplo:

$$B = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} \qquad D = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 4 \\ 3 \end{pmatrix}.$$

Observação. Sabia que um vetor no plano (ou no espaço) pode ser considerado como uma matriz coluna? Mais adiante (capítulo de Sistemas Lineares) usaremos essa forma ao representar a solução de um sistema de equações. Assim, se tivermos duas ou três incógnitas elas podem ser alocadas numa forma vetorial no plano ou no espaço, respectivamente; você também encontrará essa notação no livro "Um curso de geometria analítica e álgebra linear", citado na bibliografia comentada.

1.2.4 Matriz Nula

A matriz nula é uma matriz cujos elementos são todos nulos. Por exemplo:

$$O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad O = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Esses tipos de matrizes geralmente são denotados pela letra maiúscula O e dependendo do problema deverá discernir a ordem da matriz no exercício ou problema em questão. Alguns autores denotam essa matriz da forma: $O = \left[0_{ij} \right]_{m \times n}$.

1.2.5 Matriz Quadrada

Uma matriz quadrada é uma matriz onde o número de linhas é igual ao número de colunas. Nas seguintes matrizes, A é uma matriz de ordem n e B uma matriz de ordem 3:

Para facilitar, usamos apenas a notação $A = \left[a_{ij} \right]_n$ para representar, de forma abreviada, matrizes quadradas de ordem n.

$$A = \left[a_{ij} \right]_n \qquad B = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & 2 \\ 3 & 7 & 0 \end{pmatrix}$$

No caso de matrizes quadradas, é possível definir duas diagonais:

- A **diagonal principal** de uma matriz quadrada está dada pelos elementos na posição $i = j$. Por exemplo, os valores 1, -1 e 0 são os elementos da diagonal principal da matriz B.
- A **diagonal secundária** está dada pelos elementos da matriz cujos índices contabilizam o valor $i+j=n+1$, assim, na mesma matriz B dada acima os elementos 1, -1 e 3 são aqueles cujos índices sempre somam $i+j=3+1=4$, esses elementos são b_{13}, b_{22} e b_{31}.

Exemplo 3: Considere a matriz $B=\begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 2 \\ 3 & 7 & 0 \end{bmatrix}$.

Os elementos {1, -1, 0} formam a diagonal principal e os elementos {3, -1, 1} formam a diagonal secundária.

A partir de agora, falaremos um pouco mais sobre matrizes quadradas.

1.2.6 Matriz Diagonal

A matriz diagonal é uma matriz quadrada cujos elementos fora da diagonal principal são nulos, isto é, $a_{ij}=0$ se $i \neq j$. Por exemplo:

$$D=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix} \qquad E=\begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}.$$

Pelo fato das matrizes diagonais possuírem elementos, quase sempre não nulos. Apenas na posição (i,i) é que elas podem ser denotadas como $diag\{d_1,d_2,\ldots,d_n\}$, ou ainda na forma $diag\{d_1,d_2,\ldots,d_n\}$ onde $d_1,d_2,\ldots,d_n$ indicam os elementos diagonais. Por exemplo, a matriz D dada anteriormente pode ser escrita como $D=diag\{1,3,6\}$.

1.2.7 Matriz Identidade

A matriz identidade é uma matriz diagonal onde todos os elementos da diagonal principal são iguais a um. É geralmente denotada

com a letra I e com um índice que denota a ordem, como ilustrado a seguir:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

1.2.8 Matriz Triangular Superior

A matriz triangular superior é uma matriz quadrada de ordem n cujos elementos a_{ij} são nulos quando $i > j$. Isto é:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}.$$

1.2.9 Matriz Triangular Inferior

A matriz triangular inferior é uma matriz quadrada de ordem n cujos elementos a_{ij} são nulos quando $i < j$, ou seja:

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

1.2.10 Matriz Simétrica

Uma matriz quadrada S, de ordem n, é simétrica se $a_{ij} = a_{ji}$, para quaisquer valores dos índices i, j. São exemplos de matrizes simétricas:

$$S_2 = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \qquad S_4 = \begin{pmatrix} 1 & 0 & -1 & 3 \\ 0 & 1 & 4 & -5 \\ -1 & 4 & 0 & 0 \\ 3 & -5 & 0 & a \end{pmatrix}.$$

Quando falamos de elementos assumindo qualquer valor real podemos denotá-los com $a \in \mathbb{R}$. Nesse caso, o símbolo $\in$ é lido como "pertence a" e $\mathbb{R}$ denota os números reais.

Observe que o elemento a na posição $(4,4)$ da matriz S_4 não tem valor numérico, isto é, assume qualquer valor real.

Exemplo 4. Encontre os valores de t, w, s, z, a, b para obtermos S simétrica:

$$S = \begin{bmatrix} a & 2 & 0 & -t \\ x & b & w & 0 \\ z & -z & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Solução. Pela definição de matriz simétrica, todos os elementos s_{ij} da matriz S devem ser tais que $s_{ij} = s_{ji}$. Como a matriz é de ordem $n = 4$ e considerando que i, j variam entre 1 e 4 (ou seja, $i, j = 1, \ldots, 4$), encontramos que:

$$s_{21} = x = 2 = s_{12}.$$

Também:

$$s_{31} = z = 0 = s_{13},$$

e de forma similar:

$$s_{41} = 1 = -t = s_{14}.$$

Assim,

$$t = -1.$$

Também,

$$s_{32} = -z = w = s_{23},$$

como $z = 0$ e o oposto de zero é ele próprio, então:

$$w = 0.$$

Por último,

$$s_{11} = a \quad \text{e} \quad s_{22} = b,$$

mas não há nenhuma condição para esses valores. Portanto, a e b são valores reais quaisquer, isto é, $a, b \in \mathbb{R}$.

1.2.11 Matriz Anti-simétrica

Uma matriz quadrada A é anti-simétrica se $a_{ij} = -a_{ji}$. São exemplos de matrizes anti-simétricas as matrizes:

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 2 & 6 \\ -2 & 0 & -4 \\ -6 & 4 & 0 \end{bmatrix}$$

Exemplo 5. Considere a matriz S fornecida no **Exemplo 3**; encontre os valores de t, w, s, z, a, b para S ser uma matriz anti-simétrica.

Solução. Usando um raciocínio similar ao usado no Exemplo 3 e considerando que para cada valor de i e j deve se satisfazer $a_{ij} = -a_{ji}$, encontra-se $x = -2$, $z = 0$, $t = 1$, $w = 0$, $a = 0$ e $b = 0$. Assim:

$$S = \begin{bmatrix} 0 & 2 & 0 & -1 \\ -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Você percebeu que os elementos da diagonal principal das matrizes anti-simétricas fornecidas são todos nulos? Isso seria apenas uma coincidência? No exemplo seguinte, provaremos que esse resultado vale para qualquer matriz anti-simétrica.

Exemplo 6. Prove que os valores da diagonal principal de uma matriz anti-simétrica qualquer são todos nulos.

Solução. Se $A = \left[a_{ij} \right]_n$ é uma matriz anti-simétrica de ordem n, os seus elementos satisfazem a relação $a_{ij} = -a_{ji}$ para quaisquer valores i, j.

Os elementos na diagonal principal encontram-se na posição $i = j$, então $a_{ii} = -a_{ii}$.

Daí, $2a_{ii} = 0$ para qualquer valor de i. Em consequência, $a_{ii} = 0$ para qualquer i.

Um exemplo numérico que ilustra o que acabamos de provar foi dado no **Exemplo 4**. Nele, você encontrou que os valores diagonais são todos nulos!

1.2.12 Matriz Elementar

Uma matriz é denominada elementar se for obtida por meio de uma única mudança na matriz identidade. Essa mudança pode ser de um dos seguintes tipos:

1) A troca de uma linha (ou coluna) por outra linha (ou coluna);

2) A multiplicação de uma linha (ou coluna) por um valor $\alpha \in \mathbb{R}$;

3) A soma de uma linha (ou coluna), multiplicada pelo valor $\alpha \in \mathbb{R}$, com outra linha (ou coluna).

Exemplos:

a) A matriz elementar de ordem 2 obtida ao trocarmos a linha 1 pela linha 2 da matriz identidade de ordem 2 é dada por:

$$E_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

b) A matriz elementar de ordem 4 obtida ao multiplicar na linha 3 da matriz identidade (de ordem 4) por -2 é dada por:

$$E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

c) A matriz elementar de ordem 3 obtida ao multiplicar a linha 3 por -3 e somar com a linha 2 da matriz identidade (de ordem 3) é dada por:

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -3 \end{bmatrix}$$

Também, são matrizes elementares as matrizes:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \qquad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Agora é com você!

Exercício 3. Como foram obtidas as matrizes elementares A e B anteriores?

1.2.13 Igualdade de Matrizes

Duas matrizes A e B, de ordem $m \times n$, são ditas serem iguais se todos os seus elementos são iguais. Isso pode ser expressado com a seguinte relação de igualdade:

O símbolo matemático $\forall$ é lido "para todo". Na relação dada, $\forall$ *i, j* é lido "para todo *i* e para todo *j*".

$$a_{ij} = b_{ij}, \quad \forall i, j.$$

A expressão

$$a_{ij} = b_{ij}, \quad \forall i, j$$

também pode ser colocada como:

$$a_{ij} = b_{ij}, \ \forall i \in \{1, \ldots, m\}, \forall j \in \{1, \ldots, n\}.$$

Exemplo 7. Forneça condições para estabelecer a igualdade das matrizes A e S dadas a seguir.

$$A = \begin{bmatrix} 0 & 2 & 0 & -1 \\ -2 & 0 & 2 & -t \\ 0 & -2 & 0 & 0 \\ 1 & t & 0 & 0 \end{bmatrix} \qquad S = \begin{bmatrix} s_{11} & 2 & 0 & -1 \\ -2 & y & 2 & -t \\ 0 & -2 & 0 & 0 \\ 1 & -t & 0 & 0 \end{bmatrix}$$

Solução. Como as matrizes são de ordem 4, teremos $i, j \in \{1, \ldots, 4\}$. Se $A = S$, então, $a_{ij} = s_{ij} \ \forall i, j \in \{1, \ldots, 4\}$, assim:

$$a_{11} = 0 = s_{11}$$

$$a_{22} = 0 = s_{22} = y,$$

daí resulta:

$$y = 0.$$

Também,

$$a_{24} = -t = s_{24} = -t,$$

com isso:

$$t \in \mathbb{R}.$$

Mais,

$$a_{42} = t = s_{42} = -t$$

$$2t = 0,$$

que implica:

$$t = 0.$$

Por último, como $t \in \mathbb{R}$ e $t = 0$, implica $t = 0$.

Observação. As matrizes $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ possuem os mesmos elementos, mas não são iguais, você pode justificar o porquê?

Agora é com você!

Exercício 4. Quais são os valores de b para a matriz

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ b & 1 & b & 1 \\ 1 & b & 1 & 1 \\ 1 & 1 & b & b \end{bmatrix} \text{ ser simétrica?}$$

1.3 Operações com Matrizes

A seguir, serão definidas as operações de adição, produto por um escalar e produto de matrizes.

1.3.1 Adição de Matrizes

Dadas as matrizes $A = \left[a_{ij}\right]_{m \times n}$ e $B = \left[b_{ij}\right]_{m \times n}$, a adição das matrizes A e B é a matriz $C = \left[c_{ij}\right]_{m \times n}$, onde $c_{ij} = a_{ij} + b_{ij}$, $\forall i, j$.

Notação. $C = A + B$.

$$A + B = \left[a_{ij} + b_{ij}\right]_{m \times n}$$

Exemplo 8. Se $A = \begin{bmatrix} 0 & 2 & 0 & -1 \\ -2 & 0 & 2 & -t \\ 0 & -2 & 0 & 0 \\ 1 & t & 0 & 0 \end{bmatrix}$ e $S = \begin{bmatrix} s_{11} & 2 & 0 & -1 \\ 2 & y & -2 & -t \\ 0 & -2 & 0 & 0 \\ 1 & -t & 0 & 0 \end{bmatrix}$, calcule $C = A + S$ para t, y e s_{11} quaisquer números reais.

Solução. Ao aplicarmos a definição de soma de matrizes nas matrizes A e S, teremos:

$$C = \begin{bmatrix} s_{11} & 4 & 0 & -2 \\ 0 & y & 0 & -2t \\ 0 & -4 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{bmatrix}.$$

1.3.2 Produto de uma matriz por um escalar

Escalar
Na maioria dos casos, é um número real $\alpha \in \mathbb{R}$. É possível, também, tomarmos os escalares como números complexos, $\alpha \in \mathbb{C}$. Os escalares podem ser tomados de qualquer sistema numérico no qual podemos somar, subtrair, multiplicar e dividir de acordo com as leis habituais da aritmética.

Dado o escalar α, o produto da matriz A pelo escalar é uma matriz da mesma ordem cujos elementos foram multiplicados pelo valor α. Em outras palavras, se $A = \left[a_{ij}\right]_{m\times n}$ e $\alpha \in \mathbb{R}$, o produto de A pelo escalar α é uma matriz C de elementos c_{ij} com $c_{ij} = \alpha a_{ij}$ para todos os valores i, j definidos na matriz A. Isto é:

$$C = \left[c_{ij}\right]_{m\times n}, \text{ tal que } c_{ij} = \alpha a_{ij}, \ \forall i, j.$$

Notação. $C = \alpha A = \left[\alpha a_{ij}\right]_{m\times n}$.

Exemplo 9. Multiplique a matriz I_4 pelo escalar $\alpha = -2$.

Solução.

$$C = \alpha I_4 = -2\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

Nota. Quando $\alpha = -1$, podemos escrever $-1 \cdot A = -A$.

1.3.3 Produto de Matrizes

Dadas as matrizes $A = [a_{ik}]_{m\times t}$ e $B = \left[b_{kj}\right]_{t\times n}$, o produto das matrizes A e B é uma matriz $C = \left[c_{ij}\right]_{m\times n}$ cujos elementos c_{ij} são da forma:

$$c_{ij} = \sum_{k=1}^{t} a_{ik} b_{kj}.$$

Isto é, ao definirmos as matrizes

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1t} \\ a_{21} & a_{22} & \cdots & a_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mt} \end{bmatrix}, \quad B=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{t1} & b_{t2} & \cdots & b_{tn} \end{bmatrix} \text{ e } C=\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix},$$

os elementos da matriz produto adotam a forma:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{it}b_{tj}$$

$$c_{ij} = \sum_{k=1}^{t} a_{ik}b_{kj} .$$

Note que o número de colunas da matriz A é igual ao número de linhas da matriz B.

Notação: $C = AB = \left[\sum_{k=1}^{t} a_{ik}b_{kj} \right]_{3\times 4}$.

Exemplo 10. Seja a matriz $A=\begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix}$ e a matriz B, de ordem 3×4 com elementos $b_{ij} = i^j$. Obter a matriz produto $C = AB$.

Solução. Como o número de colunas de A é igual ao número de linhas de B, o produto pedido é possível. As matrizes explicitadas são dadas respectivamente por:

$$A=\begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix} \text{ e } B=\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 8 & 16 \\ 3 & 9 & 27 & 81 \end{bmatrix}.$$

Para obtermos a matriz produto $C = AB = \left[c_{ij} \right]_{3\times 4}$ com elementos

$$c_{ij} = \sum_{k=1}^{3} a_{ik}b_{kj}, \; i = 1,\ldots,3, \; j = 1,\cdots,4.$$

Percorrendo cada valor de i e j dado temos os elementos da:

- Primeira linha:

$$c_{11} = (2)(1) + (3)(2) + (4)(3) = 2 + 6 + 12 = 20$$
$$c_{12} = (2)(1) + (3)(4) + (4)(9) = 2 + 12 + 36 = 50$$
$$c_{13} = (2)(1) + (3)(8) + (4)(27) = 2 + 24 + 108 = 134$$
$$c_{14} = (2)(1) + (3)(16) + (4)(81) = 2 + 48 + 324 = 374$$

- Segunda linha:

$$c_{21} = (3)(1) + (4)(2) + (5)(3) = 3 + 8 + 15 = 26$$
$$c_{22} = (3)(1) + (4)(4) + (5)(9) = 3 + 16 + 45 = 64$$
$$c_{23} = (3)(1) + (4)(8) + (5)(27) = 3 + 32 + 135 = 170$$
$$c_{24} = (3)(1) + (4)(16) + (5)(81) = 3 + 64 + 405 = 472$$

- E, por último, os da terceira linha:

$$c_{31} = (4)(1) + (5)(2) + (6)(3) = 4 + 10 + 18 = 32;$$
$$c_{32} = (4)(1) + (5)(4) + (6)(9) = 4 + 20 + 54 = 78;$$
$$c_{33} = (4)(1) + (5)(8) + (6)(27) = 4 + 40 + 162 = 206;$$
$$c_{34} = (4)(1) + (5)(16) + (6)(81) = 4 + 80 + 486 = 570.$$

Sendo assim, temos a seguinte matriz:

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \end{bmatrix} = \begin{bmatrix} 20 & 50 & 134 & 374 \\ 26 & 64 & 170 & 472 \\ 32 & 78 & 206 & 570 \end{bmatrix}.$$

Ao multiplicarmos matrizes devemos tomar cuidado com a ordem das linhas e colunas, ou seja, poderemos fazer o produto de matrizes quando o número de colunas da primeira matriz for igual ao número de linhas da segunda. Assim, a matriz produto C terá um número de linhas igual ao número de linhas da matriz A e um número de colunas igual ao número de colunas de B.

1.3.4 Propriedades das Operações com Matrizes

Considere $A = \left[a_{ij}\right]_{m\times n}$, $B = \left[b_{ij}\right]_{m\times n}$ e $C = \left[c_{ij}\right]_{m\times n}$, então temos as seguintes propriedades:

1) Propriedades da Adição

A_1) Comutatividade: $A + B = B + A$;

A_2) Associatividade: $(A + B) + C = A + (B + C)$;

A_3) Elemento Neutro da Soma: $A + O = A$, $O = [0]_{m\times n}$;

A_4) Elemento Simétrico: $A + (-A) = O \quad (A - A = O)$.

Observação. $-A = -1 \cdot A = \left[-1 \cdot (a_{ij})\right]_{m \times n} = \left[-a_{ij}\right]_{m \times n}$

- **Prova das Propriedades**

A_1) Comutatividade: $A + B = B + A$

Seja $A = \left[a_{ij}\right]_{m \times n}$ e $B = \left[b_{ij}\right]_{m \times n}$

$$A + B = \left[a_{ij}\right]_{m \times n} + \left[b_{ij}\right]_{m \times n}$$
$$= \left[a_{ij} + b_{ij}\right]_{m \times n}$$
$$= \left[(a_{ij} + b_{ij})\right]_{m \times n}.$$

Usando a propriedade comutativa dos números reais:

$$(x + y) = (y + x), \text{ com } x, y \in \mathbb{R}$$

temos:

$$= \left[(b_{ij} + a_{ij})\right]_{m \times n}$$
$$= \left[b_{ij} + a_{ij}\right]_{m \times n}$$
$$= B + A.$$

Logo,

$$A + B = B + A.$$

A_2) Associatividade: $(A + B) + C = A + (B + C)$.

Consideremos $A = \left[a_{ij}\right]_{m \times n}$, $B = \left[b_{ij}\right]_{m \times n}$ e $C = \left[c_{ij}\right]_{m \times n}$.

Da definição de soma de matrizes,

$$A + B = \left[a_{ij} + b_{ij}\right]_{m \times n}$$
$$\text{e } \ (A + B) + C = \left[(a_{ij} + b_{ij}) + c_{ij}\right]_{m \times n}.$$

Usando a propriedade associativa dos números reais:

$$(x + y) + z = x + (y + z) \text{ com } x, y, z \in \mathbb{R}.$$

Temos, então:

$$= \left[a_{ij} + (b_{ij} + c_{ij})\right]_{m \times n}.$$

E usando a definição de soma de matrizes:

$$= \left[a_{ij}\right]_{m \times n} + \left[b_{ij} + c_{ij}\right]_{m \times n}$$

$$= A + (B + C).$$

Logo,

$$(A + B) + C = A + (B + C).$$

A$_3$) Elemento Neutro da Soma: $A + O = A,\ O = [0]_{m\times n}$.

Seja $A = \left[a_{ij}\right]_{m\times n}$ e $O = [0]_{m\times n}$

$$A + O = \left[a_{ij} + 0\right]_{m\times n} = \left[(a_{ij} + 0)\right]_{m\times n}.$$

Pela propriedade dos números reais:

$$x + 0 = x \ \text{ com } \ x \in \mathbb{R}.$$

Então,

$$a_{ij} + 0 = a_{ij},\ \forall i, j.$$

Com isso,

$$\left[a_{ij} + 0\right]_{m\times n} = \left[a_{ij}\right]_{m\times n} = A.$$

Logo,

$$A + O = A.$$

A$_4$) Elemento Simétrico: $A + (-A) = O$.

Seja $A = \left[a_{ij}\right]_{m\times n}$ e $-A = \left[-a_{ij}\right]_{m\times n}$. Logo,

$$A + (-A) = \left[a_{ij} + (-a_{ij})\right]_{m\times n}.$$

Pela propriedade dos números reais:

$$x + (-x) = 0 \ \text{ com } \ x \in \mathbb{R}.$$

Então,

$$a_{ij} + (-a_{ij}) = 0,\ \forall i, j.$$

Assim,

$$\left[a_{ij} + (-a_{ij})\right]_{m\times n} = [0]_{m\times n} = O.$$

Logo,

$$A + (-A) = O.$$

2) Propriedades do Produto por um Escalar

Sejam A e B duas matrizes da mesma ordem e α, β dois escalares, então:

M$_1$) $\alpha(\beta A) = (\alpha\beta)A;$

M$_2$) $\alpha(A + B) = \alpha A + \alpha B;$

M_3) $(\alpha+\beta)A=\alpha A+\beta A$;

M_4) $1\cdot A=A$.

Observação. Quando trabalhamos com matrizes, pode acontecer a necessidade de multiplicá-las pelo escalar zero, dando como resultado a matriz nula. Isto é, $0\cdot A=O$.

Você observou as diferenças entre o zero escalar e a matriz zero, denotada pela letra O?

Vejamos: se $A=\left[a_{ij}\right]_{m\times n}$, $O=[0]_{m\times n}$ e o escalar nulo (0):

$$\begin{aligned}O\cdot A&=0\left[a_{ij}\right]_{m\times n}\\&=\left[0\cdot a_{ij}\right]_{m\times n}\\&=[0]_{m\times n}\\&=O.\end{aligned}$$

- **Prova das Propriedades**

M_3) $(\alpha+\beta)A=\alpha A+\beta A$.

Sejam α,β dois escalares e a matriz $A=\left[a_{ij}\right]_{m\times n}$, então:

$$\begin{aligned}(\alpha+\beta)\cdot A&=(\alpha+\beta)\left[a_{ij}\right]_{m\times n}\\&=\left[(\alpha+\beta)a_{ij}\right]_{m\times n}.\end{aligned}$$

Usando a propriedade distributiva dos números reais

$$(x+y)\cdot z=xz+yz$$

para cada elemento da matriz, temos:

$$\begin{aligned}&=\left[(\alpha a_{ij})+(\beta a_{ij})\right]_{m\times n}\\&=\left[\alpha a_{ij}\right]_{m\times n}+\left[\beta a_{ij}\right]_{m\times n}.\end{aligned}$$

Pela definição de produto por um escalar,

$$\begin{aligned}&=\alpha\left[a_{ij}\right]_{m\times n}+\beta\left[a_{ij}\right]_{m\times n}\\&=\alpha A+\beta A.\end{aligned}$$

Logo,

$$(\alpha+\beta)\cdot A=\alpha A+\beta A.$$

M_4) $1\cdot A=A$.

Seja $A = \left[a_{ij} \right]_{m \times n}$ e o escalar $1 \in \mathbb{R}$.

$$1 \cdot A = 1 \cdot \left[a_{ij} \right]_{m \times n}$$
$$= \left[1 \cdot a_{ij} \right]_{m \times n}$$
$$= \left[(1 \cdot a_{ij}) \right]_{m \times n}.$$

Usando a propriedade do elemento neutro da multiplicação dos números reais,

$$1 \cdot x = x, \quad \forall x \in \mathbb{R}.$$

Temos:

$$\left[1 \cdot a_{ij} \right]_{m \times n} = \left[a_{ij} \right]_{m \times n} = A.$$

Logo,

$$1 \cdot A = A.$$

Agora é com você!

Exercício 5. Prove as outras propriedades do produto de uma matriz por um escalar.

Ao enunciar as propriedades do produto de matrizes não explicitamos a ordem das mesmas, por exemplo, em P_1, $(AB)C = A(BC)$ supomos possíveis os produtos AB e BC, isto é, o número de colunas de A é igual ao número de linhas de B e o número de colunas de B é igual ao número de linhas de C.

3) Propriedades do Produto de Matrizes

Considere A, B e C matrizes, então valem as seguintes propriedades de produto de matrizes:

P_1) Associativa: $(AB)C = A(BC)$;

P_2) Distributiva: $A(B + C) = AB + AC$;

P_3) $(A + B)C = AC + BC$;

P_4) $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

- **Prova das Propriedades**

P_3) $(A + B)C = AC + BC$;

Sejam as matrizes $A = [a_{ik}]_{m \times p}$, $B = [b_{ik}]_{m \times p}$, $C = [c_{kj}]_{p \times n}$, então:

$$(A + B)C = [(a_{ik} + b_{ik})]_{m \times p} \cdot \left[c_{kj} \right]_{p \times n}.$$

Usando a definição do produto de matrizes para $A + B$ e C, temos:

$$= \left[\sum_{k=1}^{p} (a_{ik} + b_{ik}) c_{kj} \right]_{m \times n}.$$

Usando a propriedade distributiva dos números reais:

$$=\left[\sum_{k=1}^{p} a_{ik}c_{kj}+b_{ik}c_{kj}\right]_{m\times n}.$$

Pela propriedade 2 dos somatórios e da definição de adição de matrizes,

A lista de propriedades encontra-se no final desta Seção.

$$\left[\sum_{k=1}^{p} a_{ik}c_{kj}+\sum_{k=1}^{p} b_{ik}c_{kj}\right]_{m\times n}=\left[\sum_{k=1}^{p} a_{ik}c_{kj}\right]_{m\times n}+\left[\sum_{k=1}^{p} b_{ik}c_{kj}\right]_{m\times n}.$$

Pela definição do produto de matrizes:

$$=AC+BC.$$

Logo,

$$(A+B)C=AC+BC.$$

P_4) $\alpha(AB)=(\alpha A)B=A(\alpha B).$

Seja $\alpha\in\mathbb{R}$, $A=[a_{ik}]_{m\times t}$ e $B=[b_{kj}]_{t\times n}$

$$\alpha(AB)=\alpha[a_{ik}]_{m\times t}\cdot[b_{kj}]_{t\times n}=\left[\alpha\sum_{k=1}^{t}(a_{ik}\cdot b_{kj})\right]_{m\times n}.$$

Usando a propriedade do somatório:

$$c\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}cx_i,\ c\text{: constante,}$$

temos:

$$=\left[\sum_{k=1}^{t}\alpha(a_{ik}\cdot b_{kj})\right]_{m\times n}.$$

Da propriedade associativa dos números reais:

$$(xy)\cdot z=x\cdot(yz)\ \text{ com }\ x,y,z\in\mathbb{R}.$$

Temos:

$$\left[\sum_{k=1}^{t}(\alpha a_{ik})\cdot b_{kj}\right]_{m\times n}.$$

E, pela definição de produto de matrizes e produto de uma matriz por um escalar,

$$=(\alpha A)B.$$

Logo,

$$\alpha(AB)=(\alpha A)B.$$

Observação. É importante observar que em geral $AB \neq BA$, isso será ilustrado com o seguinte exemplo.

Exemplo 11. Dadas as matrizes $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, a matriz produto $AB = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$, entretanto $BA = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$, verificando que $AB \neq BA$.

No ambiente virtual da disciplina você encontrará algumas atividades nas quais poderá praticar tanto a multiplicação de matrizes numéricas, usando problemas do cotidiano, quanto a aplicação das propriedades.

Agora é com você!

Exercício 6. Prove as outras propriedades do produto de matrizes.

1.3.5 Transposta de uma Matriz

Na literatura é também usual encontrarmos a transposta de uma matriz denotada como A^T ou A^t, mas usaremos tal notação pelo fato de ser a forma como trabalharemos computacionalmente com alguns softwares como MATLAB® ou SCILAB®, durante as nossas aulas ou no ambiente virtual.

Seja $A = \left[a_{ij}\right]_{m\times n}$, a matriz transposta de A, denotada por A', é aquela matriz obtida trocando-se as linhas pelas colunas de A. Isto é:

$$A' = \left[a_{ji}\right]_{n\times m}.$$

Por exemplo, se $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}_{2\times 3}$, a matriz transposta é uma matriz de ordem 3×2 dada por:

$$A' = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}_{3\times 2}.$$

Observe que na matriz transposta cada elemento na linha i e coluna j aparece como sendo um elemento da linha j e coluna i da matriz A.

Exemplo 12. Seja A uma matriz de ordem 2, encontre o valor de x de modo que $A' = A$.

$$A = \begin{bmatrix} 1 & x \\ -1 & 0 \end{bmatrix}.$$

Solução.

$$A' = \begin{bmatrix} 1 & -1 \\ x & 0 \end{bmatrix}.$$

Como $A' = A$ é uma condição do exercício, então:

$$\begin{bmatrix} 1 & -1 \\ x & 0 \end{bmatrix} = \begin{bmatrix} 1 & x \\ -1 & 0 \end{bmatrix}.$$

Isso será válido apenas se $x = -1$.

Observação. Outra forma de definirmos a matriz simétrica é usando a matriz transposta. Assim, diremos que uma matriz é simétrica se ela coincide com a sua transposta, isto é, $A' = A$.

1.3.6 Propriedades da Matriz Transposta

Dadas as matrizes A e B, são válidas as propriedades da matriz transposta:

1) $(A')' = A$;

2) $(A+B)' = A' + B'$;

3) $(AB)' = B'A'$;

4) $(\alpha A)' = \alpha A'$, $\alpha \in \mathbb{R}$.

- **Prova da Propriedade 3**

$$(AB)' = B'A'.$$

Sejam $A = [a_{ik}]_{m \times p}$, $B = [b_{kj}]_{p \times n}$

$$AB = \left[\sum_{k=1}^{p} a_{ik} b_{kj} \right]_{m \times n}$$

$$= [c_{ij}]_{m \times n}.$$

Assim:

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}.$$

Pela definição de transposta de uma matriz,

$$(AB)' = [c_{ji}]_{n\times m}$$
$$= \left[\sum_{k=1}^{p} a_{jk} b_{ki} \right]_{n\times m} \quad (1)$$

Pode-se verificar que:

$$\sum_{k=1}^{p} b_{jk} a_{ki} = \sum_{k=1}^{p} a_{jk} b_{ki}. \quad (2)$$

Por outro lado:

$$B' = [b_{jk}]_{n\times p}, \quad A' = [a_{ki}]_{p\times m}.$$

Observe que $k \in \{1,\ldots,p\}$, e

$$B'A' = \left[\sum_{k=1}^{p} b_{jk} a_{ki} \right]_{n\times m}$$

(deixamos a você a tarefa de pesquisar a propriedade do somatório usado), substituindo (2) e (1):

$$(AB)' = \left[\sum_{k=1}^{p} b_{jk} a_{ki} \right]_{n\times m}.$$

Logo,

$$(AB)' = B'A'.$$

Agora é com você!

Exercício 7. Prove as demais propriedades, justificando todos os passos do seu procedimento.

Exercício 8. Prove que se $A' = -A$, então A é anti-simétrica.

Exercício 9. Dado um escalar não nulo α, prove que, se A é uma matriz simétrica e B é uma matriz anti-simétrica, então, $\frac{1}{\alpha}A$ é simétrica e $\frac{1}{\alpha}B$ é anti-simétrica.

Exemplo 13. Prove que toda matriz quadrada pode ser colocada como a soma de uma matriz simétrica com outra anti-simétrica.

Solução. Seja $A=\left[a_{ij}\right]_n$. Em primeiro lugar, vejamos que $A+A'$ é uma matriz simétrica.

Seja $A=\left[a_{ij}\right]_n$ e $A=B+C$ com B simétrica e C anti-simétrica (ambas de ordem n). Isto é, $B'=B$ e $C'=-C$.

Transpondo, $A'=B'+C'$.

Somando a última expressão na equação $A=B+C$, temos:

$$A'+A=(B'+B)+(C'+C).$$

Sendo B simétrica e C anti-simétrica:

$$A'+A=2B.$$

Então,

$$B=\frac{A+A'}{2}.$$

Como C é anti-simétrica, ao substituirmos as equações:

$$\begin{aligned} A-A' &= B+C-B'-C' \\ &= B-B'+(C-C') \\ &= 2C \end{aligned}$$

Então:

$$C=\frac{A-A'}{2}.$$

Assim,

$$A=\left(\frac{A+A'}{2}\right)+\left(\frac{A-A'}{2}\right).$$

1.3.7 Potência de uma Matriz: A^p

Seja A uma matriz quadrada e p um número inteiro positivo, a potência p da matriz A, denotada por A^p está definida por:

$$A^p=\underbrace{A\cdot\ldots\cdot A}_{p\ \text{vezes}}$$

Exemplo 14. Seja $A = [a_{ij}]_n$, com $a_{ij} = i - j$, calcule A^3, para $p = 2, 3, 4$.

Solução. Pela lei de formação fornecida obtemos facilmente o valor de A:

Se $n = 2$,

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Assim,

$$A^2 = AA = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Se $p = 3$,

$$A^3 = AAA = A^2A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Deixamos como exercício calcular A^4.

Observações:

1) Calcular A^p equivale a calcular $A^{p-1} \cdot A$. Assim, se quiser encontrar A^{50}, calcule A^{49} e multiplique o resultado por A (para o que previamente calculou o valor de A^{48} e assim por diante).

2) Por definição, se $p = 0$ e $A \neq O$, então $A^0 = I$.

1.3.8 Traço de uma Matriz

Dada $A = \left[a_{ij}\right]_n$, o traço de A, denotado por $Tr(A)$, é o número dado pela soma dos elementos da diagonal principal. Isto é:

$$Tr(A) = \sum_{i=1}^{n} a_{ii} \;.$$

Por exemplo, se

$$A = \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 0 & 2 & 5 \\ 3 & 4 & 7 & 1 \\ 0 & 0 & 0 & 5 \end{bmatrix} \Rightarrow Tr(A) = 1 + 0 + 7 + 5 = 13.$$

1.3.9 Propriedades do Traço

Dados $A=\left[a_{ij}\right]_n$ e $B=\left[b_{ij}\right]_n$, são verdadeiras as seguintes propriedades:

1) $Tr(A+B)=Tr(A)+Tr(B)$;

2) $Tr(\alpha A)=\alpha Tr(A)$;

3) $Tr(A')=Tr(A)$;

4) $Tr(AB)=Tr(BA)$.

- **Prova da Propriedade 1**

$$Tr(A+B)=Tr(A)+Tr(B)$$

Sejam $A=\left[a_{ij}\right]_n$ e $B=\left[b_{ij}\right]_n$ duas matrizes quadradas.

Pela definição do traço,

$$Tr(A+B)=\sum_{i=1}^{n}(a_{ii}+b_{ii}),$$

e pela propriedade do somatório:

$$=\sum_{i=1}^{n}a_{ii}+\sum_{i=1}^{n}b_{ii}$$

$$=Tr(A)+Tr(B).$$

Agora é com você!

Exercício 10. Prove as outras propriedades.

Exercícios Resolvidos

1) Dada a matriz $A=\begin{bmatrix}1 & -1 & 7\\ 0 & 5 & -2\end{bmatrix}_{2\times 3}$, encontre a sua transposta.

Solução.

$$A'=\begin{bmatrix}1 & 0\\ -1 & 5\\ 7 & -2\end{bmatrix}_{3\times 2}.$$

2) Encontre o traço de matriz identidade.

Solução. Seja I_n a matriz identidade de ordem n.

$$Tr(I_n) = \sum_{i=1}^{n} 1 = n.$$

3) Encontre o traço de uma matriz diagonal e de uma matriz triangular de qualquer ordem.

Solução. Usando a notação simplificada, temos a matriz diagonal $D = diag\{d_1, d_2, \ldots, d_n\}$. Assim:

$$Tr(D) = \sum_{i=1}^{n} d_i.$$

Deixamos para você o cálculo do traço no caso de se ter uma matriz triangular.

1.3.10 Propriedades de Somatórios

Os seguintes itens fornecem algumas propriedades de somatórios úteis para a prova das propriedades listadas anteriormente.

a) $\sum_{i=1}^{n} b_i = \sum_{j=1}^{n} b_j;$

b) $\sum_{i=1}^{n} (a_i + b_i) = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i;$

c) $\sum_{i=1}^{n} b_i a_k = a_k \sum_{i=1}^{n} b_i;$

d) $\sum_{i=1}^{n} \sum_{j=1}^{m} b_{ij} = \sum_{j=1}^{m} \sum_{i=1}^{n} b_{ij}.$

Observação. No final deste Capítulo você encontrará um resumo de todas as propriedades até aqui utilizadas, que servirá de ajuda ao resolver exercícios de demonstração.

Agora é com você!

Exercício 11. Dadas as matrizes:

$$A = \begin{bmatrix} 1 & 2 & -1 \\ -2 & -3 & 2 \\ 1 & 4 & 5 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 & 3 \\ 2 & -1 & 4 \\ -3 & -1 & -17 \end{bmatrix},$$

encontre:

a) $C = A + 2B$;

b) $C = B^2$;

c) $tr(A)$, $tr(B)$ e $tr(AB)$;

Expresse as matrizes A e B como somas de uma matriz simétrica com outra anti-simétrica.

Exercício 12. Sejam as matrizes A e B, de ordem 4, $A = \left[a_{ij}\right]_4$ com $a_{ij} = \begin{cases} j^i & \text{se } i \geq j \\ 0 & \text{se } i < j \end{cases}$, e B uma matriz simétrica com $b_{ij} = i + j$ se $i \leq j$. Encontre:

a) $C = 2A - 3B$.

b) $C = B^2$. C é uma matriz simétrica?

Exercício 13. Sejam A e B matrizes simétricas, justifique se os enunciados a seguir são falsos ou verdadeiros:

- $A + B$ é uma matriz simétrica.
- AB é uma matriz simétrica.

Nota. Se sua resposta for verdade, prove. Se for falsa, apresente um contraexemplo.

Exercício 14. Imagine uma situação cotidiana e procure problematizá-la de tal forma que você possa fazer uso:

- da soma de matrizes;

- da subtração de matrizes;
- do produto de matrizes.

1.4 Determinantes

1.4.1 Menor de uma Matriz: M_{ij}

Dada uma matriz quadrada, $A=[a_{ij}]_n$, o menor da matriz A, denotado por M_{ij}, é uma submatriz de ordem $(n-1)$ obtida ao cancelarmos a linha i e a coluna j.

Assim, se:

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix} \Rightarrow M_{ij}=\left[a_{ij}\right]_{(n-1)}.$$

Com:

$$M_{ij}=\begin{bmatrix} a_{11} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{(i-1)1} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)n} \\ a_{(i+1)1} & \cdots & a_{(i+1)(j-1)} & a_{(i+1)(j+1)} & \cdots & a_{(i+1)n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n(j-1)} & a_{n(j+1)} & \cdots & a_{nn} \end{bmatrix}.$$

Exemplo:

$$\text{Se} \quad A=\begin{bmatrix} 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 1 & 3 & 4 \\ 2 & -1 & 3 & 2 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix},$$

então, o menor M_{34} é obtido ao eliminarmos a linha 3 e a coluna 4, isto é:

$$M_{34} = \begin{bmatrix} 2 & 3 & 4 & 6 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Similarmente, ao eliminarmos a linha 1 e a coluna 1, obtemos o menor M_{11}.

$$M_{11} = \begin{bmatrix} 0 & 1 & 3 & 4 \\ -1 & 3 & 2 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Agora é com você!

Exercício 15. Verifique que $A = \left[a_{ij}\right]_n$ (com n^2 elementos) possui n^2 menores.

Nessa parte da teoria assumimos que você está familiarizado(a) com o cálculo de determinantes de matrizes de ordem 2 e 3. O valor do determinante de uma matriz A é denotado nas formas $\det(A)$, $\det A$ ou $|A|$. Por exemplo, se:

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \text{ então } \det(A) = (0)(0) - (-1)(1) = 1.$$

Similarmente, se:

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \text{ então:}$$

$$\det(B) = (1)(5)(9) + (2)(6)(7) + (3)(4)(8) - (3)(5)(7) - (1)(6)(8) - (2)(4)(9) =$$
$$= 45 + 84 + 96 - 105 - 48 - 72 = 0.$$

Com esses exemplos, estamos relembrando de forma rápida que o determinante de uma matriz de ordem 2 é calculado de uma única maneira: o produto dos elementos da diagonal principal menos o produto dos elementos da diagonal secundária. E o determinante de uma matriz de ordem 3 é calculado pela Regra de Sarrus.

Para lembrar esta regra pesquise na Internet ou em algum material de matemática do ensino médio.

1.4.2 Cofator de uma Matriz: A_{ij}

O cofator A_{ij} do elemento na posição (i, j) de uma matriz A é dado pelo valor do determinante M_{ij}, multiplicado pelo valor $(-1)^{i+j}$. Isto é:

$$A_{ij} = (-1)^{i+j}\det(M_{ij}).$$

Ou:

$$A_{ij} = (-1)^{i+j}|M_{ij}|.$$

Exemplo 15. Se $A = \begin{bmatrix} 2 & 3 & 4 & 6 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$, calcule A_{44}, A_{11}, A_{31}, A_{33}, A_{14}, A_{23} e A_{32}.

Solução.

$$A_{44} = (-1)^{4+4}|M_{44}| = (+1)\begin{vmatrix} 2 & 3 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix} = 0$$

$$A_{11} = (-1)^{1+1}|M_{11}| = (+1)\begin{vmatrix} 0 & 1 & 4 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{vmatrix} = 0$$

$$A_{31} = (-1)^{3+1}|M_{31}| = (+1)\begin{vmatrix} 3 & 4 & 6 \\ 0 & 1 & 4 \\ 1 & 1 & 1 \end{vmatrix} = 19 - 18 = 1$$

$$A_{33} = (-1)^{3+3}|M_{33}| = (+1)\begin{vmatrix} 2 & 3 & 6 \\ 0 & 0 & 4 \\ 1 & 1 & 1 \end{vmatrix} = 12 - 8 = 4$$

$$A_{14} = (-1)^{1+4}|M_{14}| = (-1)\begin{vmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{vmatrix} = 0$$

$$A_{23} = (-1)^{2+3}|M_{23}| = (-1)\begin{vmatrix} 2 & 3 & 6 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{vmatrix} = 0$$

$$A_{32} = (-1)^{3+2}\left|M_{32}\right| = (-1)\begin{vmatrix} 2 & 4 & 6 \\ 0 & 1 & 4 \\ 1 & 1 & 1 \end{vmatrix} = -(18-14) = -4$$

Observe as mudanças de sinais dos elementos nas posições (i, j), isto é, $(-1)^{i+j}$:

$$\begin{matrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{matrix}$$

Em geral, para uma matriz de qualquer ordem, as mudanças de sinais dos elementos nas posições (i, j) $\left((-1)^{i+j}\right)$ são:

$$\begin{matrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{matrix}$$

1.4.3 Determinante de A usando Cofatores

Dada A uma matriz de ordem n, $A = \left[a_{ij}\right]_n$.

Se $n = 2$, os menores e os cofatores da linha um da matriz de ordem dois são dados respectivamente por:

$$M_{11} = [a_{22}],\ A_{11} = a_{22},$$

$$M_{12} = [a_{21}],\ A_{12} = -a_{21}.$$

E o valor do determinante será:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\,a_{22} - a_{21}\,a_{12}$$

$$= a_{11}\,|M_{11}| + a_{12}(-|M_{12}|)$$

$$= a_{11}A_{11} + a_{12}A_{12}.$$

Se $n = 3$, o valor do determinante da matriz (colocado em função dos cofatores relativos à primeira linha) será:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$$

$$= a_{11}(a_{22}\,a_{33} - a_{32}\,a_{23}) + a_{12}(a_{23}\,a_{31} - a_{33}\,a_{21}) + a_{13}(a_{21}\,a_{32} - a_{31}\,a_{22})$$

$$= a_{11}(+|M_{11}|) + a_{12}(-|M_{12}|) + a_{13}(+|M_{13}|)$$

$$= a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} = \sum_{j=1}^{3} a_{1j}A_{1j}.$$

Note que calculamos o determinante de A usando cofator onde $i = 1$. Podemos usar qualquer linha da matriz. Por exemplo, com $i = 2$:

$$|A| = a_{21}(-|M_{21}|) + a_{22}(+|M_{22}|) + a_{23}(-|M_{23}|)$$

$$= a_{21}A_{21} + a_{22}A_{22} + a_{23}A_{23} = \sum_{j=1}^{3} a_{2j}A_{2j}$$

Por exemplo, se $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ o determinante usando a segunda linha é dado por:

$$|A| = 4A_{21} + 5A_{22} + 6A_{23}$$

$$= 4(-|M_{21}|) + 5(+|M_{22}|) + 6(-|M_{23}|)$$

$$= -4(18 - 24) + 5(9 - 21) - 6(8 - 14)$$

$$= 24 - 60 + 36 = 0.$$

No caso geral de uma matriz de ordem n, o cálculo do determinante da matriz referido à linha 1 (ou a qualquer linha k) é dado por:

$$|A| = a_{11}\,A_{11} + a_{12}\,A_{12} + \ldots + a_{1n}\,A_{1n}$$

$$= \sum_{i=1}^{n} a_{1i}A_{1i} = \sum_{j=1}^{n} a_{1j}A_{1j}\ .$$

Se o desenvolvimento do determinante for referido a qualquer linha k, temos:

$$|A| = \sum_{j=1}^{n} a_{kj}A_{kj},$$

onde k é um valor fixo.

Por exemplo, na matriz do **Exemplo 14**, calculamos o determinante pelo desenvolvimento de cofatores referido à linha 3 (pois tendo todos seus elementos nulos evitaremos cálculos desnecessários). Assim:

$$|A| = 0 \cdot A_{31} + 0 \cdot A_{32} + 0 \cdot A_{33} + 0 \cdot A_{34} = 0.$$

Nota. Fica como regra: ao calcular o determinante usando cofatores, escolha a linha (ou coluna) da matriz que tiver o maior número de elementos nulos.

Similarmente, é possível fazer o desenvolvimento por colunas. Veja:

1) Usando a primeira coluna:

$$|A| = a_{11}A_{11} + a_{21}A_{21} + \cdots + a_{n1}A_{n1}.$$

2) Deixamos para você chegar ao seguinte desenvolvimento para uma coluna k qualquer:

$$|A| = \sum_{i=1}^{n} a_{ik}A_{ik}.$$

O desenvolvimento dado acima para encontrarmos o valor do determinante (usando linhas ou colunas) é comumente conhecido como o desenvolvimento de Laplace.

Astrônomo e matemático francês, Marquês de Pierre Simon de Laplace (1749-1827) ficou conhecido como o "Newton francês". Sua carreira foi importante por suas contribuições técnicas para as ciências exatas, tanto pelo ponto de vista filosófico que ele desenvolveu durante sua vida, quanto pela parcela que tomou parte na formação das modernas disciplinas científicas.

1.4.4 Definição Geral do Determinante de uma Matriz

- **Permutação**

Dados os n números (ou n objetos distintos) uma permutação desses números (ou objetos) consiste em dispô-los em uma determinada ordem.

Exemplo 16. Considere os números 1, 2 e 3, podemos ordená-los de várias formas, por exemplo: (1 2 3), (3 2 1), etc.

O mesmo acontece quando escolhemos 4 números, como 1, 2, 3 e 4. Podemos ordená-los, por exemplo: (1 2 3 4), (2 1 3 4), etc.

Notação. Uma permutação de n números é denotada por $(j_1 \ j_2 \ \ldots \ j_n)$.

• Número de Permutações

Dados os números 1 e 2 há duas permutações, (1 2) e (2 1), ou seja, 2! permutações.

No caso dos números 1, 2 e 3 as permutações (1 2 3) e (3 2 1) são dois exemplos, no total existem 3! permutações. Quais são?

Dado n números, $1, 2, \ldots, n$, existem $n!$ permutações.

Agora é com você!

Exercício 16. Calcule o número de permutações possíveis de 4 números.

• Inversão

É o número de mudanças necessárias em uma permutação para voltá-la à sua posição ordenada inicial.

Notação. Uma inversão de n números será denotada por:

$$J = J\,(j_1\, j_2 \ldots j_n).$$

Por exemplo, nas permutações dadas acima:

$$J\,(1\ 2\ 3) = 0,\ J\,(1\ 2\ 3\ 4) = 0 \text{ e } J\,(3\ 2\ 1) = 3.$$

No último caso, embora o número 2 esteja na posição que lhe corresponde, para colocarmos os números 3 e 1 nos seus lugares será necessário fazermos assim:

$$(3\ 2\ 1) \to (2\ 3\ 1) \to (2\ 1\ 3) \text{ e por último } (1\ 2\ 3).$$

$$\text{Ou } (3\ 2\ 1) \to (3\ 1\ 2) \to (1\ 3\ 2) \text{ e por último } (1\ 2\ 3).$$

Em ambos os casos haverá 3 inversões.

Exemplo 17. Construir uma tabela do número de inversões possíveis de 2 e 3 números.

Solução. Se $n = 2$, considere os números 1 e 2.

Permutação	N° de inversões
12	$J\,(1\ 2) = 0$
21	$J\,(2\,1) = 1$

Se $n = 3$, considere os números 1, 2 e 3.

Permutação	N° de inversões
123	$J(1\ 2\ 3) = 0$
132	$J(1\ 3\ 2) = 1$
213	$J(2\ 1\ 3) = 1$
231	$J(2\ 3\ 1) = 2$
312	$J(3\ 1\ 2) = 2$
321	$J(3\ 2\ 1) = 3$

Agora é com você!

Exercício 17. Verifique que o número de inversões da permutação J (4 3 2 1) é igual a 6.

Exemplo 18. Construir uma tabela do número de inversões de 4 números.

Solução. Neste caso o número de inversões para cada permutação $(j_1\ \ j_2\ \ j_3\ \ j_4)$ será dado por $J = J(j_1\ \ j_2\ \ j_3\ \ j_4)$. O resultado será colocado na segunda coluna da tabela.

Permutação	N° de inversões
1234	0
1243	1
1324	1
1342	2
1432	3
1423	2
2134	1
2143	:
2314	:
2341	:
2431	:
:	:

Deixamos para você completar a tabela.

• Determinante

Definição. Dada a matriz de ordem n, $A=\left[a_{ij}\right]_n$, o determinante de A é definido por:

$$\det(A)=\sum_{\rho}(-1)^J a_{1j_1}a_{2j_2}\ldots a_{nj_n}.$$

Onde $J=J(j_1\,j_2\ldots j_n)$ indica o número de inversões da permutação $(j_1\,j_2\ldots j_n)$, ρ indica que o somatório é estendido a todas as $n!$ permutações dos números $1,2,\ldots,n$.

Exemplo 19. Verifique o uso da definição nos casos dos determinantes de ordem 2 e 3.

Solução. Na solução deste exemplo serão usados os resultados obtidos no **Exemplo 15.**

Se $n=2$, então $\rho=2$, assim:

$$\det(A)=\sum_{\rho}(-1)^{J(j_1\;j_2)}a_{1j_1}a_{2j_2}=(-1)^0a_{11}a_{22}+(-1)^1a_{12}a_{21}=a_{11}a_{22}-a_{12}a_{21}.$$

Se $n=3$, $\rho=6$ e, assim:

$$\det(A)=\sum_{\rho}(-1)^{J(j_1\,j_2\,j_3)}a_{1j_1}a_{2j_2}a_{3j_3}$$

$$=(-1)^0a_{11}a_{22}a_{32}+(-1)^1a_{11}a_{23}a_{32}+(-1)^1a_{12}a_{22}a_{33}+(-1)^2a_{12}a_{23}a_{31}+(-1)^2a_{13}a_{21}a_{32}+(-1)^3a_{13}a_{22}a_{31}$$

$$=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{11}a_{23}a_{33}-a_{12}a_{21}a_{33}$$

Agora é com você!

Exercício 18.

a) Obtenha o desenvolvimento para o caso de um determinante de ordem 4.

b) Verifique a relação desse desenvolvimento com o desenvolvimento dos cofatores.

• Propriedades do Determinante

Considere A e B matrizes quadradas. Então, valem as propriedades dos determinantes.

1) Se A possui uma linha (ou colunas) de zeros, então, $\det(A)=0$;

2) Se A possui duas linhas (ou colunas) iguais, então,

$$\det(A) = 0;$$

3) Se B é obtida de A multiplicando-se uma linha (ou coluna) por um escalar α, então, $\det(B) = \alpha \det(A)$;

4) Se B é obtida por troca das posições relativas de duas linhas (ou colunas) da matriz A, então, $\det(B) = -\det(A)$;

5) Se B é obtida de A, substituindo-se a linha i (ou coluna) por ela somada a um múltiplo escalar de outra linha j (ou coluna) $(j \neq i)$ então, $\det(B) = \det(A)$;

6) $\det(A) = \det(A')$;

7) $\det(AB) = \det(A)\det(B)$.

Observações. Não é objetivo do presente material didático fazer as demonstrações das propriedades anteriores, porém as mesmas podem ser provadas a partir da definição do determinante.

Mais detalhes a respeito dessas demonstrações podem ser encontrados no livro de Álgebra Linear, de Callioli (1993), citado no final deste Capítulo.

Na **Seção 1.4.3**, ao calcularmos o determinante usando cofatores, usamos o desenvolvimento (referentes às linhas) dado por

$$\det(A) = \sum_{i=1}^{n} a_{ki} A_{ki},$$

onde k é a k-ésima linha escolhida.

Podemos enunciar uma oitava propriedade usando desenvolvimentos similares.

8) $\sum_{i=1}^{n} a_{ki} A_{li} = 0$, $l \neq k$, k, l valores fixos.

Verifiquemos a propriedade com o seguinte exemplo.

Se $k = 1$, $l = 2$ e $n = 2$:

$$\sum_{i=1}^{2} a_{1i} A_{2i} = a_{11} A_{21} + a_{12} A_{22}.$$

Assim, se

$$A=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix},$$

então:

$$A_{21}=-2 \text{ e } A_{22}=1,$$

dessa forma:

$$\sum_{i=1}^{2} a_{1i}A_{2i}=1(-2)+2(1)=0.$$

Também, ao usarmos o desenvolvimento pelas colunas e escolhendo $l=2$, $k=1$, encontramos também que:

$$\sum_{i=1}^{2} a_{il}A_{ik}=\sum_{i=1}^{2} a_{i2}A_{i1}=2(4)+4(-2)=0.$$

Agora é com você!

Exercício 19. Use as operações elementares e o Método de Laplace para encontrar o determinante das matrizes:

$$A=\begin{bmatrix}2 & 1 & 3 & 4\\ 3 & 2 & 4 & 5\\ 2 & 2 & 1 & 0\\ -2 & 1 & 3 & 1\end{bmatrix},\ B=\begin{bmatrix}1 & 1 & 2 & -1\\ 2 & 3 & 0 & -2\\ -1 & 1 & 2 & 2\\ 1 & 2 & 1 & 1\end{bmatrix},\ C=\begin{bmatrix}1 & 0 & 3\\ 2 & -1 & 4\\ -6 & -1 & -17\end{bmatrix}.$$

Exercício 20. Usando apenas as propriedades dos determinantes mostre que $\det(A)=\det(B)$. Das matrizes,

$$A=\begin{bmatrix}a & c+2a\\ b & d+2b\end{bmatrix},\ B=\begin{bmatrix}a & c\\ b & d\end{bmatrix}.$$

1.5 Matriz Adjunta: $Adj\,(A)$

Dada $A=\left[a_{ij}\right]_n$, a matriz adjunta de A é dada por

$$Adj\,(A)=\left(Cof\,(A)\right)',$$

onde $Cof\,(A)$ é a matriz cujos elementos são os cofatores A_{ij} da matriz A, ou seja, é a matriz onde cada elemento a_{ij} é igual ao cofator A_{ij} da matriz A. Um exemplo para essa definição é o seguinte:

Se $B=\begin{bmatrix}1 & -2\\ 2 & -4\end{bmatrix}$, então, os cofatores são:

$$A_{11}=(-1)^{1+1}\det[-4]=-4$$
$$A_{12}=(-1)^{1+2}\det[2]=-2$$
$$A_{21}=(-1)^{2+1}\det[-2]=2$$
$$A_{22}=(-1)^{2+2}\det[1]=1$$

$$Cof(B)=\begin{bmatrix}-4 & -2\\ 2 & 1\end{bmatrix}.$$

Assim,

$$Adj(B)=\begin{bmatrix}-4 & -2\\ 2 & 1\end{bmatrix}^{T}=\begin{bmatrix}-4 & 2\\ -2 & 1\end{bmatrix}.$$

Exemplo 20. Calcule a matriz adjunta de A dada por:

$$A=\begin{bmatrix}2 & 1 & 0\\ -3 & 1 & 4\\ 1 & 6 & 5\end{bmatrix}.$$

Solução. A matriz de cofatores de A é dada por:

$$Cof(A)=\begin{bmatrix}A_{11} & A_{12} & A_{13}\\ A_{21} & A_{22} & A_{23}\\ A_{31} & A_{32} & A_{33}\end{bmatrix}=\begin{bmatrix}-19 & 19 & -19\\ -5 & 10 & -11\\ 4 & -8 & 5\end{bmatrix},$$

pois:

$$A_{11}=(-1)^{1+1}\det\begin{bmatrix}1 & 4\\ 6 & 5\end{bmatrix}=5-24=-19$$

$$A_{12}=(-1)^{1+2}\det\begin{bmatrix}-3 & 4\\ 1 & 5\end{bmatrix}=-(-15-4)=19$$

$$A_{13}=(-1)^{1+3}\det\begin{bmatrix}-3 & 1\\ 1 & 6\end{bmatrix}=-18-1=-19$$

$$A_{21}=(-1)^{2+1}\det\begin{bmatrix}1 & 0\\ 6 & 5\end{bmatrix}=-(5+0)=-5$$

$$A_{22}=(-1)^{2+2}\det\begin{bmatrix}2 & 0\\ 1 & 5\end{bmatrix}=10-0=10$$

$$A_{23}=(-1)^{2+3}\det\begin{bmatrix}2 & 1\\ 1 & 6\end{bmatrix}=-(12-1)=-11$$

$$A_{31} = (-1)^{3+1}\det\begin{bmatrix} 1 & 0 \\ 1 & 4 \end{bmatrix} = 4 - 0 = 4$$

$$A_{32} = (-1)^{3+2}\det\begin{bmatrix} 2 & 0 \\ -3 & 4 \end{bmatrix} = -(8-0) = -8$$

$$A_{33} = (-1)^{3+3}\det\begin{bmatrix} 2 & 1 \\ -3 & 1 \end{bmatrix} = 2 + 3 = 5$$

Assim,

$$\begin{aligned} Adj\,(A) &= Cof\,(A)' \\ &= \begin{bmatrix} -19 & 19 & -19 \\ -5 & 10 & -11 \\ 4 & -8 & 5 \end{bmatrix}' \\ &= \begin{bmatrix} -19 & -5 & 4 \\ 19 & 10 & -8 \\ -19 & -11 & 5 \end{bmatrix}. \end{aligned}$$

Também, o determinante da matriz A é $\det(A) = -19$, pois

$$\begin{aligned} \det(A) &= a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} \\ &= 2(-19) + 1(19) + 0(-19) \\ &= -38 + 19 = -19 \end{aligned}$$

Observe que $Adj(A)A = \det(A)I_3$; considerando A do exercício anterior, temos:

$$Adj(A)A = \begin{bmatrix} -19 & -5 & 4 \\ 19 & 10 & -8 \\ -19 & -11 & 5 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 & 0 \\ -3 & 1 & 4 \\ 1 & 6 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} (-19)\cdot 2 + (-5)\cdot(-3) + 4\cdot 1 & (-19)\cdot 1 + (-5)\cdot 1 + 4\cdot 6 & (-19)\cdot 0 + (-5)\cdot 4 + 4\cdot 5 \\ 19\cdot 2 + 10\cdot(-3) + (-8)\cdot 1 & 19\cdot 1 + 10\cdot 1 + (-8)\cdot 6 & 19\cdot 0 + 10\cdot 4 + (-8)\cdot 5 \\ (-19)\cdot 2 + (-11)\cdot(-3) + 5\cdot 1 & (-19)\cdot 1 + (-11)\cdot 1 + 5\cdot 6 & (-19)\cdot 0 + (-11)\cdot 4 + 5\cdot 5 \end{bmatrix}$$

$$\begin{aligned} &= \begin{bmatrix} -19 & 0 & 0 \\ 0 & -19 & 0 \\ 0 & 0 & -19 \end{bmatrix} \\ &= -19\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \det(A)I_3 \end{aligned}$$

O próximo teorema mostra que essa afirmação é válida para qualquer matriz quadrada.

Teorema. *Se A é uma matriz de ordem n,*

$$Adj(A) \cdot A = A \cdot Adj(A) = \det(A) \cdot I_n.$$

Demonstração.

$$A \cdot Adj(A) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & \ddots & \vdots \\ A_{1n} & \cdots & A_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{j=1}^{n} a_{1j}A_{1j} & \sum_{j=1}^{n} a_{1j}A_{2j} & \sum_{j=1}^{n} a_{1j}A_{3j} & \cdots & \sum_{j=1}^{n} a_{1j}A_{(n-1)j} & \sum_{j=1}^{n} a_{1j}A_{nj} \\ \sum_{j=1}^{n} a_{2j}A_{1j} & \sum_{j=1}^{n} a_{2j}A_{2j} & \sum_{j=1}^{n} a_{2j}A_{3j} & \cdots & \sum_{j=1}^{n} a_{2j}A_{(n-1)j} & \sum_{j=1}^{n} a_{2j}A_{nj} \\ \sum_{j=1}^{n} a_{3j}A_{1j} & \sum_{j=1}^{n} a_{3j}A_{2j} & \sum_{j=1}^{n} a_{3j}A_{3j} & \cdots & \sum_{j=1}^{n} a_{3j}A_{(n-1)j} & \sum_{j=1}^{n} a_{3j}A_{nj} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \sum_{j=1}^{n} a_{(n-1)j}A_{1j} & \sum_{j=1}^{n} a_{(n-1)j}A_{2j} & \sum_{j=1}^{n} a_{(n-1)j}A_{3j} & \cdots & \sum_{j=1}^{n} a_{(n-1)j}A_{(n-1)j} & \sum_{j=1}^{n} a_{(n-1)j}A_{nj} \\ \sum_{j=1}^{n} a_{nj}A_{1j} & \sum_{j=1}^{n} a_{nj}A_{2j} & \sum_{j=1}^{n} a_{nj}A_{3j} & \cdots & \sum_{j=1}^{n} a_{nj}A_{(n-1)j} & \sum_{j=1}^{n} a_{nj}A_{nj} \end{bmatrix}$$

Usando a **Propriedade 8** dos determinantes nos elementos fora da diagonal principal, temos:

$$A \cdot Adj(A) = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}A_{1j} & 0 & 0 & \cdots & 0 \\ 0 & \sum_{j=1}^{n} a_{2j}A_{2j} & 0 & \cdots & 0 \\ 0 & 0 & \sum_{j=1}^{n} a_{3j}A_{3j} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sum_{j=1}^{n} a_{nj}A_{nj} \end{bmatrix}$$

Pelo desenvolvimento de Laplace (por linhas) temos o valor do determinante:

$$\det(A) = \sum_{j=1}^{n} a_{kj} A_{kj}, \text{ para cada } k = 1, 2, \dots, n,$$

isto é:

$$A \cdot Adj(A) = \begin{bmatrix} \det(A) & 0 & 0 & \cdots & 0 \\ 0 & \det(A) & 0 & \cdots & 0 \\ 0 & 0 & \det(A) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \det(A) \end{bmatrix},$$

$$= \det(A) I_n.$$

De forma similar, podemos encontrar

$$Adj(A) \cdot A = \det(A) \cdot I_n.$$

Assim, temos demonstrado que

$$Adj(A) \cdot A = A \cdot Adj(A) = \det(A) \cdot I_n.$$

■

1.6 Inversa de uma Matriz

1.6.1 Matriz Singular

Definição. Uma matriz é dita singular se o seu determinante é nulo. Caso contrário, dizemos que a matriz é não singular.

Por exemplo, a matriz

$$B = \begin{bmatrix} 1 & -2 \\ 2 & -4 \end{bmatrix}$$

é uma matriz singular, pois

$$\det(B) = 1 \cdot (-4) - [2 \cdot (-2)] = -4 - (-4) = -4 + 4 = 0.$$

Já a matriz identidade de ordem 3 é não singular, pois $\det(I_3) = 1$. Em geral, uma matriz identidade de ordem qualquer é não singular.

Você saberia dizer por quê? Pense a respeito!

1.6.2 Matriz Inversa

Definição. Seja A uma matriz quadrada de ordem n. Dizemos que A é inversível se existe uma única matriz B (da mesma ordem) tal que:

$$AB = BA = I_n.$$

B é denominada matriz inversa de A.

Notação. $B = A^{-1}$.

Por exemplo, se $A = \begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix}$, a matriz $B = \begin{bmatrix} -1/2 & 1/6 \\ 0 & 1/3 \end{bmatrix}$ é a respectiva matriz inversa, pois:

$$AB = BA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Propriedade. Se A é inversível, então, A é não singular.

> **Prova.** Será suficiente encontrar que o $\det(A)$ não é nulo. Demonstrando por absurdo, supomos o contrário, isto é, $\det(A) = 0$, e devemos chegar a uma contradição.

Assim, usando a **Propriedade 7** dos determinantes:

$$\begin{aligned} \det(AB) &= \det(A) \cdot \det(B) \\ &= 0 \cdot \det(B) = 0. \end{aligned}$$

Por outro lado, temos por hipótese que A é inversível, então existe B tal que $AB = I$, assim:

$$\begin{aligned} \det(AB) &= \det(I) \\ &= 1. \end{aligned}$$

Assim, $0 = 1$, impossível, é uma contradição!

Uma vez que a contradição foi encontrada, então o enunciado é verdadeiro. Assim, a propriedade fica demonstrada. Logo, A é não singular.

Geralmente uma contradição é denotada pelo símbolo $\Rightarrow\Leftarrow$.
O mesmo poderá ser usado nas próximas provas.

Conhecendo que $\det(A) \neq 0$, para A inversível, uma forma de verificar a existência da matriz inversa será encontrar o valor do determinante da matriz. Após essa verificação, o passo seguinte será encontrarmos a matriz inversa, A^{-1}. Como exemplo, nos casos das

matrizes $A=\begin{bmatrix}-2 & 1\\ 0 & 3\end{bmatrix}$ e $B=\begin{bmatrix}1 & -2\\ 2 & -4\end{bmatrix}$, podemos afirmar que apenas A possui inversa.

Como obtermos A^{-1}?

1.6.3 Cálculo da Matriz Inversa usando a Matriz Adjunta

Sabendo que existe A^{-1}, então:

$$AA^{-1} = A^{-1}A = I.$$

Observe pela propriedade da matriz adjunta que

$$A\cdot\left(\frac{Adj(A)}{\det(A)}\right)=\left(\frac{Adj(A)}{\det(A)}\right)\cdot A = I_n.$$

Assim, a única possibilidade será:

$$A^{-1}=\frac{Adj\,(A)}{\det\,(A)}.$$

Exemplo 21. Se $A=\begin{bmatrix}-2 & 1\\ 0 & 3\end{bmatrix}$, encontre A^{-1}.

Solução. Encontramos facilmente que $\det(A) = -6$, e também a matriz adjunta

$$Adj\,(A)=\begin{bmatrix}3 & 0\\ -1 & -2\end{bmatrix}.$$

Assim,

$$A^{-1}=\frac{Adj\,(A)}{\det\,(A)}=\begin{bmatrix}-1/2 & 1/6\\ 0 & 1/3\end{bmatrix}.$$

Agora é com você!

Exercício 21. Seja $A=\begin{bmatrix}2 & 1 & 0\\ -3 & 1 & 4\\ 1 & 6 & 5\end{bmatrix}$, verifique se sua matriz inversa é

$$A^{-1} = \begin{bmatrix} 1 & 5/19 & -4/19 \\ -1 & -10/19 & 8/19 \\ 1 & 11/19 & -5/19 \end{bmatrix}.$$

1.6.4 Propriedades da Inversa de uma Matriz

Se A e B são inversíveis, então:

1) $(AB)^{-1} = B^{-1}A^{-1}$;
2) $(A^{-1})^{-1} = A$;
3) $(A^t)^{-1} = (A^{-1})^t$;
4) $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

- **Prova da Propriedade 1**

$$(AB)^{-1} = B^{-1}A^{-1}$$

Em primeiro lugar, vejamos se existe $(AB)^{-1}$. Calculando $\det(AB)$:

$$\det(AB) = \det(A)\det(B).$$

Por hipótese existem as inversas das matrizes A e B $(\exists A^{-1}, \exists B^{-1})$, isto é, $\det(A) \neq 0$ e $\det(B) \neq 0$. Assim, $\det(AB) \neq 0$ e com isso $\exists\,(AB)^{-1}$, isto é,

$$(AB)(AB)^{-1} = I. \qquad (1)$$

Como:

$$A \cdot A^{-1} = I \quad \text{e} \quad B \cdot B^{-1} = I.$$

Na segunda parte dessa última relação, multiplicamos em ambos os lados pela inversa de A (pela direita):

$$(B \cdot B^{-1}) \cdot A^{-1} = I \cdot A^{-1}.$$

Associando e multiplicando por I, temos

$$B \cdot (B^{-1} \cdot A^{-1}) = A^{-1},$$

e multiplicando à esquerda por A:

$$A \cdot \big(B \cdot (B^{-1} \cdot A^{-1})\big) = A \cdot A^{-1}.$$

Você também pode considerar os seguintes passos após a expressão (2):

$$(AB)^{-1}(AB)(B^{-1}A^{-1}) = I(AB)^{-1}$$

$$I(B^{-1}A^{-1}) = I(AB)^{-1}$$

$$(B^{-1}A^{-1}) = (AB)^{-1}$$

Associando novamente e, sabendo que $AA^{-1} = I$, temos

$$(AB)(B^{-1}A^{-1}) = I. \qquad (2)$$

Sendo que a existência da matriz inversa é única e comparando as expressões (1) e (2) concluímos que

$$(AB)^{-1} = (B^{-1}A^{-1}).$$

Agora é com você!

Exercício 22. Prove as propriedades 2, 3 e 4, justificando o seu procedimento.

Ao calcular a matriz inversa de A, usando a matriz adjunta, vimos que $A^{-1} = \frac{Adj(A)}{\det(A)}$, e nos exemplos aplicamos essa relação para matrizes de ordem 2 e 3. E se a matriz for de ordem maior ou igual a 4? O procedimento acaba sendo mais trabalhoso nesses casos. Vejamos agora como podemos obter a matriz inversa sem usar a matriz adjunta.

1.6.5 Cálculo da Matriz Inversa por Operações Elementares

Seja A uma matriz não singular, portanto existe A^{-1} e $\det(A) \neq 0$. Por definição, sabemos que

$$AA^{-1} = A^{-1}A = I.$$

Então, a ideia é encontrarmos uma matriz que ao ser multiplicada por A (à direita ou à esquerda) resulte na matriz identidade. Para tal é necessário conhecer o que são aperações elementares e fazer uso das matrizes elemetares.

- **Operações Elementares**

Operações elementares são realizadas na matriz com o objetivo de invertê-la, reduzi-la ou simplesmente colocá-la num formato especificado previamente. Elas podem ser de três tipos:

1) A troca de uma linha (ou coluna) por outra linha (ou coluna);

2) A multiplicação de uma linha (ou coluna) por um valor $\alpha \in \mathbb{R}$, com $\alpha \neq 0$;

3) A soma de uma linha (ou coluna) multiplicada pelo valor $\alpha \in \mathbb{R}$, $(\alpha \neq 0)$ numa outra linha (ou coluna).

Se l_i e l_j representam a linhas i e j da matriz e α é o escalar citado anteriormente, então, as operações elementares dadas acima serão denotadas respectivamente por:

a) $l_i \leftrightarrow l_j$;

b) αl_i;

c) $l_j \leftarrow \alpha l_i + l_j$.

Seja A uma matriz, se uma (ou várias) operação elementar for efetuada nessa matriz, obteremos uma matriz diferente, a qual denotaremos por $\tilde{A}$. Assim, o processo efetuado será denotado por:

$$A \xrightarrow[elementar(es)]{operação(ções)} \tilde{A}$$

Exemplos:

Se realizarmos uma operação elementar na matriz identidade de ordem 2, I_2, e trocarmos a linha 1 pela linha 2 da matriz, obteremos a seguinte matriz elementar:

$$\tilde{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

A operação efetuada é denotada por

$$I_2 \xrightarrow{l_1 \leftrightarrow l_2} \tilde{A}.$$

Dada a matriz de ordem 4,

$$A = \begin{bmatrix} 1 & 0 & 6 & 1 \\ 0 & 1 & 0 & 0 \\ 3 & 0 & -2 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix},$$

ao fazermos a operação elementar que multiplica a linha 3 da matriz por –2, obtemos a seguinte matriz:

$$\tilde{A} = \begin{bmatrix} 1 & 0 & 6 & 1 \\ 0 & 1 & 0 & 0 \\ -6 & 0 & 4 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}.$$

Indicamos isso com:

$$A \xrightarrow{(-2)l_3} \begin{bmatrix} 1 & 0 & 6 & 1 \\ 0 & 1 & 0 & 0 \\ -6 & 0 & 4 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}.$$

Dada a matriz de ordem 3,

$$B = \begin{bmatrix} 1 & 8 & 2 \\ 0 & 1 & 0 \\ 5 & 1 & -3 \end{bmatrix},$$

ao fazermos duas operações elementares, obtemos a seguinte matriz $\tilde{B}$:

$$B \xrightarrow[3l_1]{l_2 \leftarrow (-3)l_3 + l_2} \begin{bmatrix} 3 & 24 & 6 \\ -15 & -2 & 9 \\ 5 & 1 & -3 \end{bmatrix} = \tilde{B}$$

Assim, a matriz $\tilde{B}$ foi obtida:

1) multiplicando-se a linha 3 por –3 e somando-a à linha 2 da matriz *B*,

2) multiplicando a linha 1 por 3.

Observação. A operação elementar $l_2 \leftarrow (-3)l_3 + l_2$ indica a linha onde a soma das linhas está acontecendo. No caso, a soma será efetuada na linha 2 da matriz.

Agora é com você!

Exercício 23. Dadas as matrizes $A=\begin{bmatrix}-2 & 0\\ 0 & -2\end{bmatrix}$ e $B=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$,

encontre $\tilde{A}$ e $\tilde{B}$, após as operações elementares efetuadas em A e B respectivamente. As operações são indicadas por: $A \xrightarrow[2l_1]{l_1 \leftrightarrow l_2} \tilde{A}$ e $B \xrightarrow[\substack{-5l_2 \\ l_1 \leftrightarrow l_2}]{l_4 \leftarrow 2l_3 + l_4} \tilde{B}$.

Exercício 24. Quais operações elementares devem ser feitas de modo a transformar a matriz C na sua forma triangular superior?

$$C=\begin{bmatrix}1 & 3 & 3\\ 0 & 1 & 1\\ 1 & 2 & 0\end{bmatrix}.$$

- **Forma Escada de uma Matriz**

Após efetuarmos operações elementares por linhas, na matriz inicial, dizemos que ela está na forma escada se a matriz resultante obter:

a) O primeiro elemento não nulo de uma linha não nula deve ser igual a 1;

b) A coluna que contém o primeiro elemento não nulo de alguma linha tem todos os seus outros elementos (da coluna) iguais à zero;

c) Toda linha nula ocorre abaixo de todas as linhas não nulas (isto é, daquelas que possuem pelo menos um elemento não nulo);

d) Se as linhas $1,\ldots,r$ são linhas não nulas, e se o primeiro elemento não nulo da linha i ocorre na coluna k_i (a coluna k referida à linha i), então, $k_1 < k_2 < \ldots < k_r$. (Exemplo: se $i=1$ e $k_1 = 3$, então, para $i=2$, $k_1 < k_2$ significa que k_2 será maior que 3, e assim por diante).

Exemplos:

- As seguintes matrizes encontram-se na forma escada:

$$A_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 & 5 & 0 & 3 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

- Já as seguintes matrizes não estão na forma escada:

$$B = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Agora é com você!

Exercício 25.

a) Observe as matrizes B e C. Que elementos devem ser trocados para que elas fiquem na forma escada?

b) Justifique quais condições da forma escada de uma matriz não são satisfeitas no caso das matrizes C e D.

Observações:

1) Na prática, a forma escada serve para transformar uma matriz quadrada na sua forma triangular, na qual os elementos da diagonal principal sejam uns ou zeros.

2) A prática de reduzir uma matriz usando operações elementares é um exercício muito útil para obter a inversa de uma matriz e resolver sistemas lineares.

• Operações Elementares versus Matrizes Elementares

Cada operação elementar é representada por uma matriz elementar (como definido na **Seção 1.2**) e o produto de sucessivas matrizes elementares pode nos conduzir à matriz inversa. Veremos isso com os seguintes exemplos e exercícios.

Exemplo 22. Dada a matriz A, converta-a numa matriz triangular superior.

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & -1 \\ 1 & 2 & 0 & 0 \\ -1 & 1 & 9 & 3 \end{bmatrix}$$

Solução.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & -1 \\ 1 & 2 & 0 & 0 \\ -1 & 1 & 9 & 3 \end{bmatrix} \xrightarrow[l_4 \leftarrow l_1 + l_4]{l_3 \leftarrow -l_1 + l_3} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -3 & -4 \\ 0 & 3 & 12 & 7 \end{bmatrix} \xrightarrow[\frac{l_2}{3}]{l_2 \leftrightarrow l_4} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 4 & \frac{7}{3} \\ 0 & 0 & -3 & -4 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

$$\xrightarrow{l_4 \leftrightarrow l_3} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 4 & \frac{7}{3} \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -3 & -4 \end{bmatrix} \xrightarrow[\frac{-l_4}{7}]{l_4 \leftarrow 3l_3 + l_4} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 4 & \frac{7}{3} \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \tilde{A}.$$

Assim, $\tilde{A}$ é uma matriz triangular superior.

Exemplo 23. Encontre matrizes elementares que representam as quatro primeiras operações elementares efetuadas no **Exemplo 22**.

Solução. No **Exemplo 22**, foram efetuadas sete operações elementares. Cada uma delas representará, respectivamente, as matrizes elementares $E_1, \ldots, E_7$. Assim, a primeira operação $l_3 \leftarrow -l_1 + l_3$ dá origem à matriz elementar:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Também, a operação elementar $l_4 \leftarrow l_1 + l_4$ origina a matriz elementar:

$$E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Similarmente, as operações elementares $l_2 \leftrightarrow l_4$ e $\frac{l_2}{3}$ originam as matrizes elementares:

$$E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \ E_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Agora é com você!

Exercício 26. Quais foram as matrizes E_5, E_6, E_7 no exemplo anterior?

Observe que, nas matrizes do **Exemplo 23**,

$$E_1 A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -3 & -4 \\ -1 & 1 & 9 & 3 \end{bmatrix} \text{ e } E_2(E_1 A) = E_2 \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -3 & -4 \\ -1 & 1 & 9 & 3 \end{bmatrix},$$

isto é,

$$E_2 E_1 A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -3 & -4 \\ 0 & 3 & 12 & 7 \end{bmatrix}.$$

Similarmente,

$$E_3(E_2 E_1 A) = E_3 \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -3 & -4 \\ 0 & 3 & 12 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 3 & 12 & 7 \\ 0 & 0 & -3 & -4 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

$$\text{e} \quad E_4 E_3 E_2 E_1 A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 4 & \frac{7}{3} \\ 0 & 0 & -3 & -4 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

Repetindo o processo, chegamos ao seguinte resultado:

$$E_7E_6E_5E_4E_3E_2E_1A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 4 & \frac{7}{3} \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

que é a matriz triangular superior $\tilde{A}$ obtida no **Exemplo 22**.

Agora usaremos os conceitos anteriores para aprender a calcular a matriz inversa usando operações elementares.

O **Exemplo 23** mostrou como ocorre o processo de redução de uma matriz de ordem 4 na sua forma triangular superior. Em geral esse será o processo para reduzir matrizes de ordem superior. Assim, se A é uma matriz não singular, encontraremos as matrizes elementares que transformam a matriz na forma de uma matriz identidade, isto é, encontraremos as matrizes elementares $E_1, E_2, \ldots, E_k$ tais que

$$E_k \ldots E_3E_2E_1A = I.$$

Assim, se $B = E_k \ldots E_3E_2E_1$ estaremos afirmando que $BA = I$.

Com isso, e usando a definição de matriz inversa, a inversa da matriz será dada por

$$A^{-1} = E_k \ldots E_3E_2E_1.$$

A matriz inversa não é mais do que o produto de matrizes elementares!

Ilustremos esse processo com o seguinte exemplo:

Exemplo 24. Usando matrizes elementares, vamos obter a matriz inversa da matriz A:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & -1 \\ 1 & 2 & 0 & 0 \\ -1 & 1 & 9 & 3 \end{bmatrix}.$$

Solução. Usando os resultados obtidos no **Exemplo 23** (onde se encontraram 7 matrizes elementares, $E_1,\dots,E_7$,para reduzir a matriz na forma triangular superior) obtivemos que

$$A=\begin{bmatrix}1&2&3&4\\0&0&1&-1\\1&2&0&0\\-1&1&9&3\end{bmatrix}\xrightarrow[\text{elementares}]{\text{operações}}\begin{bmatrix}1&2&3&4\\0&1&4&\frac{7}{3}\\0&0&1&-1\\0&0&0&1\end{bmatrix}.$$

Como $E_7E_6E_5E_4E_3E_2E_1A=\tilde{A}$, então

$$\det(E_7E_6E_5E_4E_3E_2E_1)\cdot\det(A)=\det(\tilde{A}).$$

Assim, $\det(A)=0$ (por quê?) e A é não singular $(\exists A^{-1})$. Continuando o processo:

$$\begin{bmatrix}1&2&3&4\\0&1&4&\frac{7}{3}\\0&0&1&-1\\0&0&0&1\end{bmatrix}\xrightarrow[l_2\leftarrow\frac{-7l_4}{3}+l_2]{l_3\leftarrow l_4+l_3}\begin{bmatrix}1&2&3&4\\0&1&4&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\xrightarrow[l_2\leftarrow-4l_3+l_2]{l_1\leftarrow-4l_4+l_1}$$

$$\begin{bmatrix}1&2&3&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\xrightarrow[l_1\leftarrow-2l_2+l_1]{l_1\leftarrow-3l_3+l_1}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}=I.$$

Serão construídas mais 6 matrizes elementares $E_8,E_9,\dots,E_{13}$, e deixamos para você a tarefa de explicitá-las!

Temos, então:

$$E_{13}E_{12}\dots E_8E_7\dots E_1A=I,$$

e, assim,

$$B=E_{13}E_{12}\dots E_8E_7\dots E_1.$$

Isto é,

$$BA=I\Rightarrow A^{-1}=B.$$

Após fazermos o produto, a matriz pedida é:

$$A^{-1} = E_{13} \ldots E_8 \ldots E_1 = \begin{bmatrix} \frac{8}{7} & \frac{18}{7} & \frac{-17}{21} & \frac{-2}{3} \\ \frac{-4}{7} & \frac{-9}{7} & \frac{19}{21} & \frac{1}{3} \\ \frac{1}{7} & \frac{4}{7} & \frac{-1}{7} & 0 \\ \frac{1}{7} & \frac{-3}{7} & \frac{-1}{7} & 0 \end{bmatrix}.$$

Observações:

- Ao tentar resolver o exemplo, você pode verificar que o conjunto de matrizes elementares encontradas no processo de escalonamento não será o único, pois dependerá da escolha das operações elementares efetuadas, não obstante a matriz inversa será a mesma.
- Fizemos questão de apresentar um exemplo com uma matriz de ordem 4, com o objetivo de facilitar a compreensão do método, além de colocar uma prática que usualmente não se expõe em livros da literatura disponível.

- **Um Método Prático**

O processo anterior foi explicado para que você entenda, passo-a-passo, como uma matriz é reduzida até ser convertida à matriz identidade. Na prática, toda vez que queiramos obter a matriz inversa de uma matriz não singular, procedemos da seguinte forma:

$$[A \vdots I] \xrightarrow[\text{elementares}]{\text{operações}} [I \vdots A^{-1}],$$

ou seja, acrescentamos à direita uma matriz identidade da mesma ordem da matriz e fazemos o processo de redução. O último resultado terá uma matriz identidade à esquerda, e à direita a inversa da matriz dada. Visto de outra forma, a matriz identidade da direita estará armazenando todas as operações elementares efetuadas no processo.

Exemplo 25. Use o método prático para obter a matriz inversa do **Exemplo 24**.

Solução.

$$\left[\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 1 & 9 & 3 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[l_3 \leftarrow l_1 + l_3]{\substack{l_2 \leftrightarrow l_4 \\ l_2 \leftarrow -l_1 + l_2}} \left[\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & 3 & 12 & 7 & 1 & 0 & 0 & 1 \\ 0 & 0 & -3 & -4 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 & 0 & 0 \end{array}\right] \xrightarrow[l_4 \leftarrow 3l_3 + l_4]{l_4 \leftrightarrow l_3}$$

$$\left[\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & 3 & 12 & 7 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -7 & -1 & 3 & 1 & 0 \end{array}\right] \xrightarrow[\substack{l_3 \leftarrow l_4 + l_3 \\ l_1 \leftarrow 4l_4 + l_1}]{\substack{l_2 \leftarrow l_4 + l_2 \\ -\frac{l_4}{7}}} \left[\begin{array}{cccc|cccc} 1 & 2 & 3 & 0 & \frac{3}{7} & \frac{12}{7} & \frac{4}{7} & 0 \\ 0 & 3 & 12 & 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 1 & 0 & \frac{1}{7} & \frac{4}{7} & \frac{-1}{7} & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{7} & \frac{-3}{7} & \frac{-1}{7} & 0 \end{array}\right] \xrightarrow[l_1 \leftarrow -3l_3 + l_1]{l_2 \leftarrow -12l_3 + l_2}$$

$$\left[\begin{array}{cccc|cccc} 1 & 2 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & \frac{-12}{7} & \frac{-27}{7} & \frac{19}{7} & 1 \\ 0 & 0 & 1 & 0 & \frac{1}{7} & \frac{4}{7} & \frac{-1}{7} & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{7} & \frac{-3}{7} & \frac{-1}{7} & 0 \end{array}\right] \xrightarrow[l_1 \leftarrow -2l_2 + l_1]{\frac{l_2}{3}} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & \frac{8}{7} & \frac{18}{7} & \frac{-17}{21} & \frac{-2}{3} \\ 0 & 1 & 0 & 0 & \frac{-4}{7} & \frac{-9}{7} & \frac{19}{21} & \frac{1}{3} \\ 0 & 0 & 1 & 0 & \frac{1}{7} & \frac{4}{7} & \frac{-1}{7} & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{7} & \frac{-3}{7} & \frac{-1}{7} & 0 \end{array}\right],$$

$$\text{assim: } A^{-1} = \begin{bmatrix} \frac{8}{7} & \frac{18}{7} & \frac{-17}{21} & \frac{-2}{3} \\ \frac{-4}{7} & \frac{-9}{7} & \frac{19}{21} & \frac{1}{3} \\ \frac{1}{7} & \frac{4}{7} & \frac{-1}{7} & 0 \\ \frac{1}{7} & \frac{-3}{7} & \frac{-1}{7} & 0 \end{bmatrix}.$$

Agora é com você!

Exercício 27. Use o método prático para encontrar as matrizes inversas:

a) da matriz $A = \begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix}$.

b) das matrizes $A = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

Exercício 28. Considerando as matrizes A, B e C, encontre a matriz inversa, se possível.

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 3 & 2 & 4 & 5 \\ 2 & 2 & 1 & 0 \\ -2 & 1 & 3 & 1 \end{bmatrix}.$$

$$B = \begin{bmatrix} 1 & 1 & 2 & -1 \\ 2 & 3 & 0 & -2 \\ -1 & 1 & 2 & 2 \\ 1 & 2 & 1 & 1 \end{bmatrix}.$$

$$C = \begin{bmatrix} 1 & 0 & 3 \\ 2 & -1 & 4 \\ -6 & -1 & -17 \end{bmatrix}.$$

Resumo

Neste Capítulo inserimos os conceitos básicos que servirão de base para o estudo dos próximos Capítulos deste Livro. Em particular, quando falamos de matrizes você, além de lembrar alguns conceitos já aprendidos no ensino médio, foi levado à reflexão e observação de propriedades antes desconhecidas, com a intenção de despertar a sua capacidade lógica e, por que não dizer, a sua capacidade de abstração, uma preparação aos conceitos que serão vistos no capítulo dos espaços vetoriais.

Você poderá encontrar também muita informação dos conteúdos aqui fornecidos (e também os conteúdos dos próximos capítulos) ao navegar por sítios com conteúdos matemáticos (confiáveis) da *internet*. Um exemplo, entre outros, que lhe pode ser útil é dado ao acessar o *link*:

<http://pt.wikibooks.org/wiki/Matem%C3%A1tica_Elementar:_Matrizes>.

Esse material de apoio, bem como algumas atividades, pode ser encontrado no ambiente virtual de aprendizagem.

Ainda, julgamos importante fornecer à você um resumo simplificado das propriedades estudadas até o momento. Ele poderá servir como material de apoio para o desenvolvimento dos exercícios de prova.

Propriedades das Matrizes

Adição de Matrizes

Sejam $A = \left[a_{ij}\right]_{m\times n}$, $B = \left[b_{ij}\right]_{m\times n}$ e $C = \left[c_{ij}\right]_{m\times n}$, então:

- Comutatividade: $A + B = B + A$;
- Associatividade: $A + (B + C) = (A + B) + C$;
- Elemento neutro: $A + O = A$;
- Elemento oposto: $A + (-A) = O$.

Produto de uma Matriz por um Escalar

Dados $A=\left[a_{ij}\right]_{m\times n}$, $B=\left[b_{ij}\right]_{m\times n}$ e $\alpha\in\mathbb{R}$ (um escalar real), então:

- Distributividade: $\alpha(A+B)=\alpha A+\alpha B$;
- Distributividade: $(\alpha+\beta)A=\alpha A+\beta A$;
- Associatividade: $\alpha(\beta A)=(\alpha\beta)A$;
- $0\cdot A=O$.

Produto de Matrizes

Dados $A=\left[a_{ik}\right]_{m\times p}$, $B=\left[b_{kj}\right]_{p\times n}$ e $C=\left[c_{jl}\right]$ (ordem conveniente), então:

- $A(BC)=(AB)C$;
- $A(B+C)=AB+AC$;
- $(A+B)C=AC+BC$;
- $\alpha(AB)=(\alpha A)B=A(\alpha B)$, $\alpha\in\mathbb{R}$. (Um escalar real).

Observe que em geral $AB\neq BA$. Também, $AB=O$, não implica $A=O$ ou $B=O$.

Transposta de uma Matriz

Dados $A=\left[a_{ij}\right]_{m\times n}$ e $B=\left[b_{ij}\right]_{m\times n}$, então:

- $(A')'=A$;
- $(A+B)'=A'+B'$
- se $B=\left[b_{ij}\right]_{n\times p}$;
- $(AB)'=B'A'$;
- $(\alpha A)'=\alpha A'$, $\alpha\in\mathbb{R}$.

Traço de uma Matriz

- $Tr(A+B)=Tr(A)+Tr(B)$;
- $Tr(\alpha A)=\alpha Tr(A)$;
- $Tr(A)=Tr(A')$;
- $Tr(AB)=Tr(BA)$.

Inversa de uma matriz

- $(AB)^{-1} = B^{-1}A^{-1}$;
- $(A^{-1})^{-1} = A$;
- $(A^{t})^{-1} = (A^{-1})^{t}$;
- $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

Bibliografia Comentada

BOLDRINI, José et al. *Álgebra linear*. 3. ed. São Paulo: Harbra, 1980.

Essa referência auxiliará para complementar e estender alguns conceitos não apresentados, nos quais você possa estar interessado(a) em conhecer e se aprofundar.

CALLIOLI, C. A. et al. *Álgebra linear e aplicações*. 6. ed. [S.l.]: Atual Editora, 1993.

Nesse livro, além de contar com os conteúdos da Álgebra Linear e Aplicações, você encontrará as provas das propriedades do determinante de uma matriz usando a definição.

KÜHLKAMP, Nilo. *Matrizes e sistemas de equações lineares*. Florianópolis: Ed. UFSC, 2005.

Esse livro possui múltiplos exemplos e problemas práticos que você pode resolver e, assim, acrescentar à sua prática. Também, são apresentados algumas aplicações e problemas de matrizes e sistemas lineares.

SANTOS, Reginaldo. *Um curso de geometria analítica e álgebra linear*. Belo Horizonte: Ed. UFMG, 2001.

Essa referência é muito útil para aplicar os conceitos usando alguns recursos computacionais, assim como os detalhes de algumas provas úteis de seu interesse.

Capítulo 2

Sistemas Lineares

Capítulo 2
Sistemas Lineares

Neste capítulo, você estudará conceitos e métodos de resolução de sistemas lineares que serão usados nos próximos capítulos. Para tal, como motivação lembraremos de algumas questões da Geometria Analítica que nos ajudarão na visualização de certos casos que conduziram à caracterização de sistemas lineares no plano e no espaço. Também, consideramos que você conheça os métodos convencionais de resolução de duas ou três equações com duas ou três incógnitas e a aplicação ***a regra de Cramer*** *nesses casos. A partir disso, levamos ao questionamento de o que fazer em casos onde esses métodos não são aplicáveis.*

A regra de Cramer é um teorema em álgebra linear que dá a solução de um sistema de equações lineares em termos de determinantes. Recebe esse nome em homenagem a Gabriel Cramer (1704-1752).

2.1 Preliminares

Antes de iniciarmos, lembraremos de alguns conceitos de geometria analítica que serão importantes para entender melhor o conteúdo que será dado neste capítulo.

Uma vez que retas e planos são subconjuntos de $\mathbb{R}^3$, forneceremos as notações das retas dadas pelos eixos coordenados e dos planos dados pelos planos coordenados.

Lembremos da notação de um ponto P qualquer do espaço com coordenadas x, y e z:

$$P = P(x, y, z) \in \mathbb{R}^3.$$

Já a notação de um vetor $\overline{a}$ com componentes a_1, a_2, a_3 é dada por:

$$\overline{a} = (a_1, a_2, a_3).$$

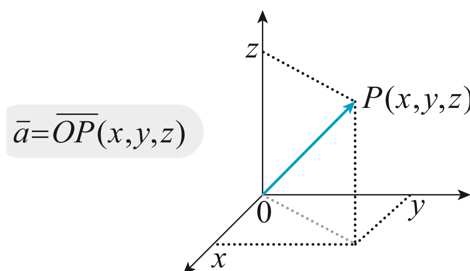


Figura 2.1. O ponto P e o vetor $\overline{a}$.

2.1.1 Eixos coordenados

Os eixos coordenados são denotados na seguinte forma vetorial:

- Eixo x

$$£_x = \{x \in \mathbb{R}^3 \,/\, x = \alpha(1,0,0),\ \alpha \in \mathbb{R}\}$$

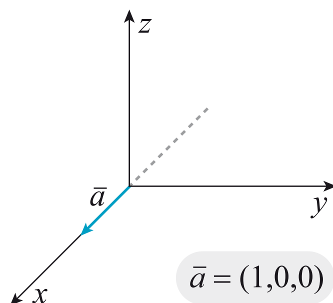


Figura 2.2. O eixo coordenado x.

- Eixo y

$$£_y = \{y \in \mathbb{R}^3 \,/\, y = \alpha(0,1,0),\ \alpha \in \mathbb{R}\}$$

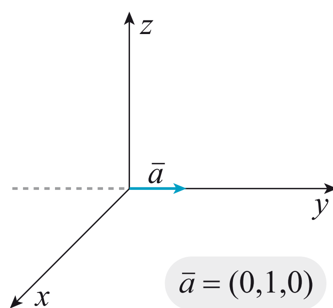


Figura 2.3. O eixo coordenado y.

- Eixo z

$$£_z = \{z \in \mathbb{R}^3 \,/\, z = \alpha(0,0,1),\ \alpha \in \mathbb{R}\}$$

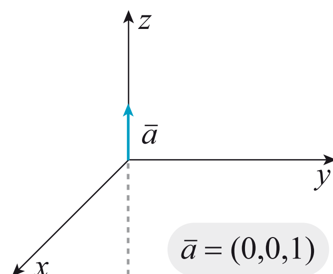


Figura 2.4. O eixo coordenado z.

2.1.2 Planos Coordenados

As formas vetoriais e cartesianas dos planos coordenados são dadas, respectivamente, da seguinte maneira:

- Plano XY

$$\Pi_{xy} = \{x \in \mathbb{R}^3 \,/\, x = \alpha(1,0,0) + \beta(0,1,0),\ \alpha, \beta \in \mathbb{R}\}$$

$$\text{ou} \quad \Pi_{xy} = \{(x, y, z) \,/\, z = 0\}$$

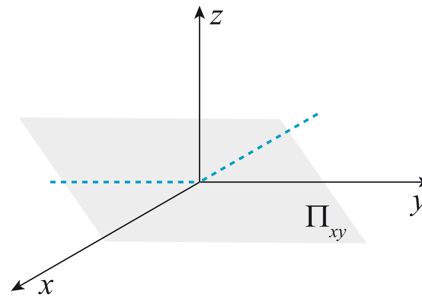


Figura 2.5. O plano Π_{xy}.

- Plano YZ

$$\Pi_{yz} = \{x \in \mathbb{R}^3 \,/\, x = \alpha(0,1,0) + \beta(0,0,1),\ \alpha, \beta \in \mathbb{R}\}$$

$$\text{ou} \quad \Pi_{yz} = \{(x, y, z) \,/\, x = 0\}$$

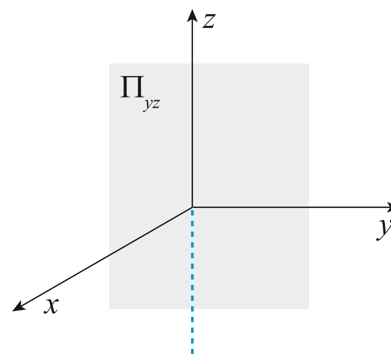


Figura 2.6. O plano Π_{yz}.

- Plano XZ

$$\Pi_{xz} = \{x \in \mathbb{R}^3 \,/\, x = \alpha(1,0,0) + \beta(0,0,1),\ \alpha, \beta \in \mathbb{R}\}$$

$$\text{ou} \quad \Pi_{xz} = \{(x,y,z) \,/\, y = 0\}$$

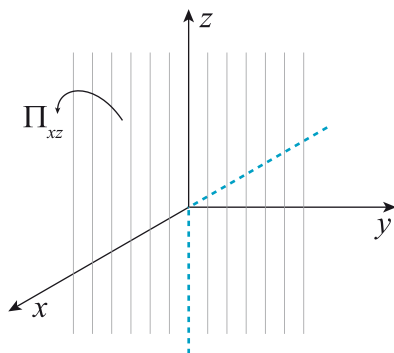


Figura 2.7. O plano Π_{xz}.

No espaço, seja ε uma reta qualquer que passa pelo ponto $P_0(x_0, y_0, z_0)$ com vetor diretor $\overline{a} = (a_1, a_2, a_3)$. Então, se $t \in \mathbb{R}$ é um parâmetro, um ponto $P(x,y,z)$ de ε é dado por:

O vetor diretor, como o nome já diz, determina a direção da reta.

$$x = x_0 + ta_1;$$

$$y = y_0 + ta_2;$$

$$z = z_0 + ta_3.$$

Essa equação é conhecida como a equação paramétrica da reta.

São equações equivalentes à equação geral da reta, da forma $x = f(t)$, $y = g(t)$ e $z = h(t)$, que relacionam as coordenadas x, y e z dos pontos da reta com um parâmetro t.
Lembre dos conteúdos da disciplina de Geometria Analítica.

Exemplo 1. Seja ε uma reta que passa pelo ponto $(1,1,2)$ e vetor diretor $\overline{a} = (1,-1,0)$. Encontre uma forma de equacionar ε que não envolva o parâmetro t.

Solução. A equação paramétrica de ε é:

$$x = 1 + t \quad \Rightarrow \quad t = x - 1$$

$$y = 1 - t \quad \Rightarrow \quad y = 1 - (x-1) = 2 - x$$

$$z = 2 + 0t = 2 \quad \Rightarrow \quad z = 2$$

Assim, $t \in \mathbb{R}$ implica $x \in \mathbb{R}$. Isto é, a reta pode ser equacionada na forma:

$$\varepsilon : \begin{cases} y = 2 - x \\ z = 2 \end{cases}$$

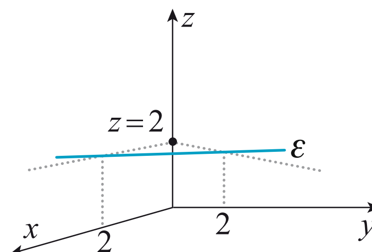


Figura 2.8. Gráfico da reta ε do **Exemplo 1.**

A equação cartesiana (ou geral) do plano Π que passa pelo ponto $P_0(x_0, y_0, z_0)$ e vetor normal $\overline{n} = (a, b, c)$ é dada por:

$$\Pi : ax + by + cz = d$$

Exemplo 2. Dado o tetraedro que passa pela origem e pelos pontos $(0,0,1)$, $(1,0,0)$ e $(0,1,0)$ como indicado na Figura 2.9, considere a face desse tetraedro no octante $X^+ \ Y^+ \ Z^+$. Então, o plano determinado por essa face tem equação:

$$\Pi : x + y + z = 1$$

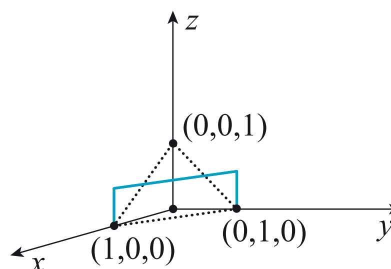


Figura 2.9. Plano definido pela face de tetraedro com vértice na origem.

Agora é sua vez!

Verifique se você está acompanhando as discussões que fizemos. Resolva o exercício que deixamos a seguir:

Exercício 1. Encontre dois vetores do plano Π, mostrado na Figura 2.9 do **Exemplo 2**, e verifique que a equação é dada por $x + y + z = 1$.

Exemplo 3. A reta ε do **Exemplo 1** pode ser considerada como sendo a intersecção do plano $\Pi_1 : x + y = 2$ e o plano $\Pi_2 : z = 2$, paralelo ao plano cartesiano Π_{xy}, conforme ilustrado na figura.

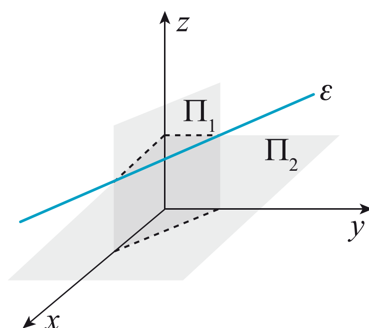


Figura 2.10. A reta ε obtida como interseção de dois planos.

Com o desenvolvimento desta seção e antes de iniciarmos a formulação algébrica dos sistemas lineares, colocaremos as várias possibilidades de posições relativas entre retas e planos e os respectivos conjuntos solução. Logo depois, veremos a conexão desta seção com a formulação algébrica a ser fornecida neste capítulo.

2.2 Sistemas Lineares

Retas em $\mathbb{R}^2$

Em $\mathbb{R}^2$, uma reta com os seus infinitos pares de pontos (x, y) é equacionada na forma

$$ax + by = c \tag{1}$$

As posições relativas entre **duas retas no plano** são:

a) **Concorrentes**, isto é, a intersecção é um ponto. Assim, dadas duas equações, existe uma única solução;

b) **Coincidentes**, isto é, as retas são iguais. Isso equivale a dizer que há muitas soluções (ou, mais apropriadamente falando, infinitas!) ou, ao resolver duas equações, existem infinitas soluções;

c) **Paralelas**, isto é, não há intersecção entre elas. Quer dizer que a solução dada por duas equações é o conjunto vazio (ou não existe solução).

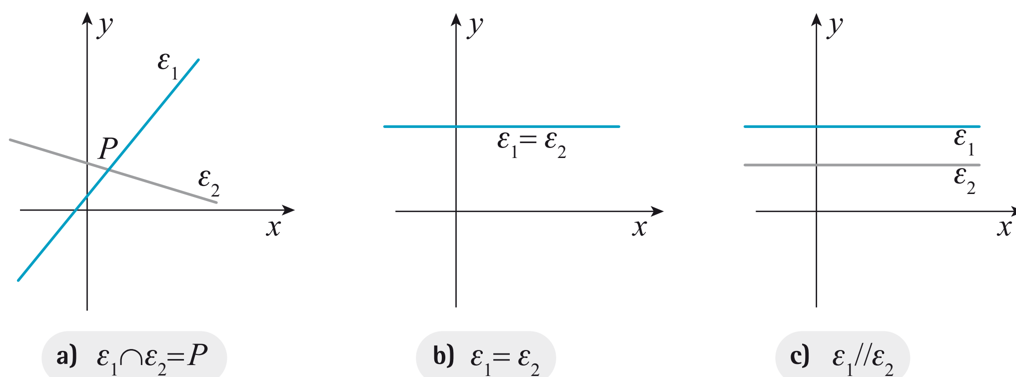


Figura 2.11. Posições relativas de retas no plano.

De forma similar, vejamos o que pode acontecer no caso de termos planos em $\mathbb{R}^3$.

Planos em $\mathbb{R}^3$

- **Um plano**

Dado um plano Π, os seus infinitos pontos (x, y, z) satisfazem a equação linear:

$$ax + by + cz = d \qquad (2)$$

Dito de outra forma, o conjunto de pontos de um plano qualquer é dado pela solução de uma equação linear com três incógnitas.

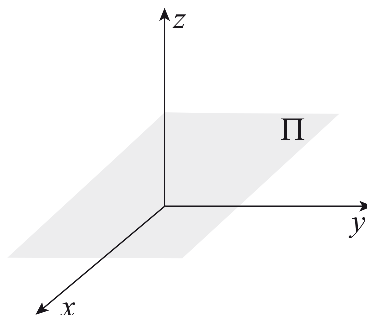


Figura 2.12. O plano Π no espaço.

- Dois planos

As posições relativas entre dois planos, dados por duas equações lineares, são as seguintes:

a) **Coincidentes**, se cada ponto $P(x,y,z)$ do conjunto solução satisfaz ambas as equações, então, existirão infinitas soluções;

b) **Paralelos**, isto é, não existem pontos $P(x,y,z)$ que satisfaçam simultaneamente as equações de cada um dos planos, isto é, não existem soluções;

c) **Concorrentes numa reta.** Neste caso, a única possibilidade é que a interseção dos planos seja uma reta de pontos $P(x,y,z)$ que satisfaz ambas as equações. Dessa forma, haverá infinitas soluções.

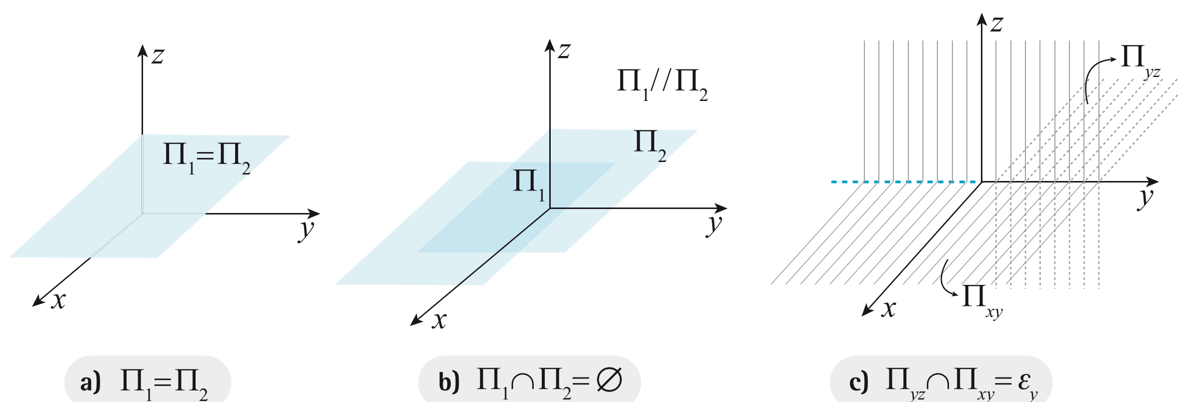


Figura 2.13. Posições relativas de dois planos no espaço.

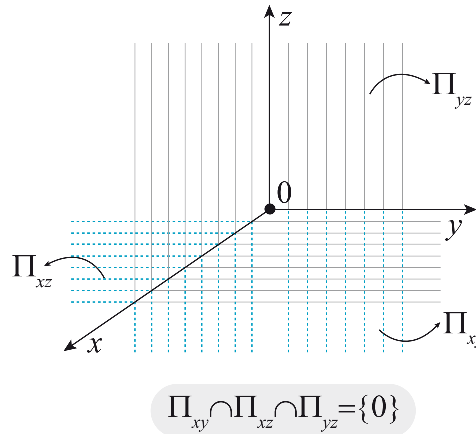
- Três Planos

As posições relativas entre três planos são as seguintes:

a) **Concorrentes num ponto.** Existe uma única possibilidade para que as equações dos planos sejam satisfeitas simultaneamente, isto é, existe uma única solução. Observe na Figura 2.14 a intersecção dada pelos três planos coordenados. Uma forma prática de visualizar essa situação é olharmos para uma esquina qualquer de uma sala de aula, na qual as paredes são consideradas como três pedaços de planos;

b) **Interseção vazia.** Não há pontos (x,y,z) satisfazendo simultaneamente as equações dos planos e é dito que a interseção é vazia; assim, não existe solução. Nessa situação, um exemplo

prático é olhar três faces quaisquer de um tijolo (cuja forma é um paralelepípedo), como indicado na Figura 2.15.



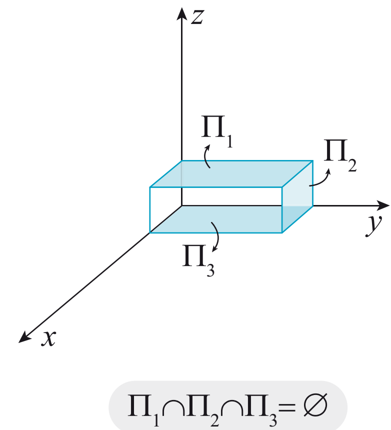


Figura 2.14. Planos concorrentes num ponto.

Figura 2.15. Planos não concorrentes.

c) **Concorrentes numa reta**, isto é, a interseção dos planos é uma reta; portanto, existem infinitas soluções. Nesse caso, um exemplo prático é observarmos um livro aberto em que as folhas representam pedaços de planos que se interceptam na lombada (e claro, teremos que imaginar uma lombada muito fina!). Assim, a lombada seria um pedaço da reta interseção.

No ambiente virtual serão exibidas essas e outras situações práticas do nosso cotidiano que ilustram as posições relativas de retas e planos.

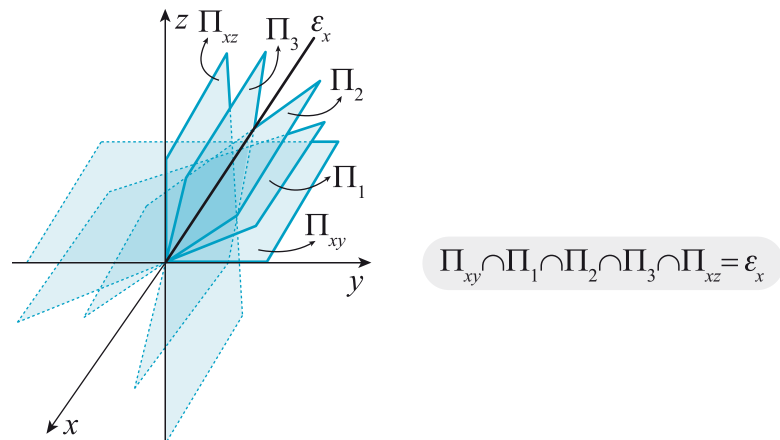


Figura 2.16. Planos concorrentes numa reta.

Em cada uma das situações dadas anteriormente, procura-se o conjunto solução de equações com duas ou três incógnitas (1) e (2). Com o objetivo de generalizar, denotamos (x_1, x_2) as incógnitas da equação (1) e (x_1, x_2, x_3) as incógnitas da equação (2). Nesses casos, escrevemos:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2, \end{aligned}$$

no caso da interseção de duas retas em $\mathbb{R}^2$. Ou, na forma:

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & = & b_2 \end{array}$$

no caso da interseção de dois planos.

Dadas m equações com n incógnitas, um sistema linear, em geral, é apresentado na seguinte forma:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned} \qquad (\mathrm{SL_g}),$$

em que $\mathrm{SL_g}$ indica um Sistema Linear geral.

Assim, se

$$A_{m\times n} = [a_{ij}]_{m\times n},\ X_{n\times 1} = [x_{ij}]_{n\times 1} \text{ e } B_{n\times 1} = [b_{ij}]_{n\times 1},$$

o sistema dado pode ser representado matricialmente por

$$A_{m\times n}X_{n\times 1} = B_{m\times 1},$$

ou, simplesmente,

$$AX = B \qquad (\mathrm{SL_m}),$$

em que $\mathrm{SL_m}$ indica um Sistema Linear na forma matricial.

A matriz A é conhecida como matriz de coeficientes, X é a matriz das incógnitas e B, a matriz dos termos independentes.

Observação. A ordem da matriz de coeficientes define a ordem do sistema. No caso, o sistema $AX = B$ é de ordem $m \times n$.

São exemplos de sistemas lineares:

i) $9x - y = 2,$
$x + 4y = 0$

ii) $2x - y + 3z = 1,$

iii) $x + y + z = 1$
$x - z = 0$
$y + 7z = 2$

onde (i) é um sistema de duas equações com duas incógnitas e (ii) é um sistema de uma equação com três incógnitas. Já o sistema (iii) possui três equações e três incógnitas.

Agora é com você!

Exercício 2. Com o intuito de verificar se você lembra do conteúdo da Seção anterior, forneça uma interpretação geométrica dos sistemas lineares dados nos exemplos (i), (ii) e (iii).

Você lembra de ter estudado a solução de sistemas de ordem dois ou três (e com duas ou três incógnitas), no Ensino Médio? Nesses casos os métodos convencionais ou a regra de Cramer sempre eram possíveis de serem aplicados. Existem sistemas onde isso pode não acontecer, por exemplo:

Exemplo 4. Resolver o seguinte sistema linear de equações de ordem 3:

$$\begin{aligned} 3x - y + z &= 4 \\ -x + 7y - 2z &= 1 \\ 2x + 6y - z &= 5 \end{aligned}$$

Solução. Usaremos métodos convencionais para resolver o sistema.

Enumerando cada equação:

$$\begin{aligned} 3x - y + z &= 4 \qquad (1) \\ -x + 7y - 2z &= 1 \qquad (2) \\ 2x + 6y - z &= 5 \qquad (3) \end{aligned}$$

Isolando z de (3), $z = 2x + 6y - 5$, e substituindo em (1) e (2), obtemos:

$$\begin{aligned} 3x - y + 2x + 6y - 5 &= 4 \\ -x + 7y - 2(2x + 6y - 5) &= 1 \end{aligned} \quad \Rightarrow \quad \begin{aligned} 5x + 5y &= 9 \\ -5x - 5y &= -9 \end{aligned} \qquad (4)$$

Na última expressão, a primeira equação é igual à segunda, com exceção do sinal. Então, qual é a solução de (4), onde $5x + 5y = 9$? Os métodos convencionais nesse caso não fornecem uma resposta.

Também, se $A = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 7 & -2 \\ 2 & 6 & -1 \end{bmatrix}$, observamos que $\det(A) = 0$, então o método de Cramer não pode ser usado! **Você sabe por quê?**

Uma interpretação geométrica do sistema dado corresponde a termos três planos se intersectando. A equação (4) junto à (3) fazem-nos obter a seguinte resposta:

$$\begin{aligned} 5x+5y&=9 \\ z&=2x+6y-5 \end{aligned} \quad \text{ou} \quad \begin{aligned} 5x+5y&=9 \\ 2x+6y-z&=5 \end{aligned} \qquad (5)$$

A expressão (5) corresponde a uma reta onde o parâmetro t não aparece explicitamente. Por exemplo, ao considerarmos $z=t$ e uma manipulação algébrica simples no sistema (5), poderemos obter os valores x,y,z em função do parâmetro t, da seguinte forma:

$$\begin{aligned} x&=\frac{29}{4}-\frac{5}{4}t \\ y&=\frac{7}{20}+\frac{1}{4}t, \quad t\in\mathbb{R} \\ z&=0+t \end{aligned}$$

Essa forma é similar ao que foi trabalhado no **Exemplo 1**, não é? E, lembrando a Figura 2.16, estamos numa situação em que temos infinitas soluções!

Observe que a resposta do **Exemplo 4** foi fornecida com uma caracterização da solução e os conceitos aprendidos na Geometria Analítica, mas... como podemos resolver esse exemplo analiticamente?

Exemplos como o anterior, em que os métodos convencionais ou a regra de Cramer não podem ser usados, serão o foco principal para o desenvolvimento dos conceitos nas próximas Seções. Uma forma de contornarmos situações como as mostradas anteriormente nos conduzirá ao uso do escalonamento de matrizes. Assim, esses sistemas serão caracterizados e as soluções serão facilmente encontradas.

2.2.1 Posto e Nulidade de uma Matriz

Posto de uma Matriz

Definição. Dada uma matriz A de ordem $m\times n$, o posto da matriz é dado pela ordem da maior submatriz não singular da matriz dada.

Notação: p(A)

Exemplo 5. Encontre o posto da matriz

$$A=\begin{bmatrix}1 & 3 & 5\\ 0 & -1 & 2\\ 0 & 2 & -4\end{bmatrix}.$$

Solução. Vejamos as submatrizes que podemos encontrar nesse caso.

Se $n=1$,

$[a_{21}]=0$ e há outras não nulas.

Se $n=2$,

escolhemos $A_1=\begin{bmatrix}3 & 5\\ -1 & 2\end{bmatrix}$ e $A_2=\begin{bmatrix}0 & -1\\ 0 & 2\end{bmatrix}$ dentre outras possíveis.

Calculamos agora os respectivos determinantes:

$$\det\left(\begin{bmatrix}3 & 5\\ -1 & 2\end{bmatrix}\right)=11,\ \det\left(\begin{bmatrix}0 & -1\\ 0 & 2\end{bmatrix}\right)=0.$$

Se $n=3$, nesse caso a única submatriz é a matriz dada com

$$\det(A)=0.$$

Dessa forma, 2 é a maior ordem da submatriz não singular, isto é, o posto de A é igual a 2.

Observação. O método não é prático para matrizes de ordem maior, pois nesse caso teremos que calcular determinantes de matrizes de ordem maior que 3 ou 4. Por exemplo, se A é de ordem 7 com determinante nulo, estaremos obrigados a calcular determinantes de matrizes de ordem 5 ou 6! Logo depois de definir a nulidade de uma matriz, apresentaremos uma forma prática de calcular o posto de uma matriz.

Nulidade de uma Matriz

Definição. Dada uma matriz A de ordem $m\times n$, a nulidade da matriz é dada pela diferença entre o número de colunas e o seu posto.

Notação: nul(A)

$$\text{nul}(A)=n-p.$$

Exemplo 6. Encontrar a nulidade da matriz dada no **Exemplo 5**.

Solução. No **Exemplo 5**, obtivemos que $\rho(A) = 2$ e, como $n = 3$, temos:

$$\text{nul}(A) = 3 - 2 = 1.$$

Observações:

- O posto e a nulidade de uma matriz são utilizados na resolução de sistemas lineares.
- O posto e a nulidade de uma matriz estão associados à dimensão do espaço linha e o espaço coluna da mesma. Esses conceitos serão vistos posteriormente.

As linhas não nulas obtidas ao reduzir uma matriz na sua forma escada são ditas linhas linearmente independentes.

O conceito de independência linear será estudado no próximo capítulo.

Teorema. *Se A é uma matriz de ordem $m \times n$, o posto da matriz é determinado pelo número de linhas linearmente independentes da matriz.*

Por exemplo, sejam A, B e C três matrizes e $\tilde{A}$, $\tilde{B}$ e $\tilde{C}$ suas respectivas formas reduzidas:

$$\tilde{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{e} \quad \tilde{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix},$$

linearmente independentes. Então, a matriz A só tem uma linha linearmente independente, a matriz B tem duas linhas linearmente independentes, e a matriz C três linhas linearmente independentes.

Assim, os postos das matrizes A, B e C são respectivamente 1, 2 e 3.

Exemplo 7. Use o teorema anterior para encontrar o posto das matrizes A e B:

a) $A = \begin{bmatrix} 1 & 3 & 5 \\ 0 & -1 & 2 \\ 0 & 2 & -4 \end{bmatrix}_{3\times 3}$, b) $B = \begin{bmatrix} 1 & 2 & -2 & 0 & -1 \\ -1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 5 & -4 & -1 & -2 \end{bmatrix}_{4\times 5}$.

Solução.

a) $A = \begin{bmatrix} 1 & 3 & 5 \\ 0 & -1 & 2 \\ 0 & 2 & -4 \end{bmatrix}$; após escalonamento, obtemos a matriz

$$\tilde{A} = \begin{bmatrix} 1 & 3 & 5 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Essa matriz tem as duas primeiras linhas linearmente independentes.

Assim, $p(A) = 2$ e, em consequência, $\text{nul}(A) = 1$.

b) Similarmente, no caso da matriz B, após operações elementares, chegamos à forma:

$$\tilde{B} = \begin{bmatrix} 1 & 2 & -2 & 0 & -1 \\ 0 & 3 & -2 & -1 & -1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Assim,

$$p(B) = 3 \text{ e } \text{nul}(B) = 5 - 3 = 2.$$

Agora é sua vez!

Exercício 3. Considerando a matriz B do **Exemplo 7**, use operações elementares para reduzi-la e obtenha o posto e a nulidade. Observe que os valores do posto e da nulidade não mudam ao obtermos matrizes reduzidas diferentes.

Propriedade. Se A é uma matriz de ordem $m \times n$, então

$$p(A) \leq \text{mín}\{m, n\}.$$

Para exibir esse resultado, usaremos como exemplo a matriz B do **Exemplo 7**, em que $m = 4$, $n = 5$, assim,

$$p(B) \leq \text{mín}\{4, 5\} = 4 \quad \Rightarrow \quad p(B) \leq 4.$$

De fato, $p(A) = 3$.

Observação. Usando a propriedade, observe que:

- Se $m < n$, então $p(A) \leq m$.
- Se $m > n$, então $p(A) \leq n$.

Como tarefa, faça uma pesquisa sobre a prova dessa propriedade.

2.2.2 Matrizes Equivalentes e Sistemas Equivalentes

Matrizes Equivalentes

Definição. Duas matrizes A e $\tilde{A}$ são ditas matrizes equivalentes se uma delas é obtida ao fazermos operações elementares na outra.

Assim, se $A \xrightarrow[\text{elementares}]{\text{operações}} \tilde{A}$, dizemos que as matrizes A e $\tilde{A}$ são equivalentes. Por exemplo, se

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix}, \quad \tilde{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{e}$$

$$B = \begin{bmatrix} 3 & 24 & 6 \\ -15 & -2 & 9 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} 1 & 8 & 2 \\ 0 & 118 & 39 \end{bmatrix},$$

então, as duplas de matrizes $A, \tilde{A}$ e $B, \tilde{B}$ são equivalentes, pois $\tilde{A}$ e $\tilde{B}$ foram obtidas após um processo de operações elementares.

Agora é sua vez!

Exercício 4. Quais operações elementares foram realizadas nas matrizes A e B para obtermos as matrizes $\tilde{A}$ e $\tilde{B}$, citadas anteriormente?

Propriedade. Matrizes equivalentes possuem o mesmo posto.

Um modo de exemplificar essa última propriedade é observarmos que uma matriz elementar, E, é sempre equivalente à matriz identi-

dade, I. **Saberia dizer por quê?** Sabendo-se que a matriz identidade possui n linhas linearmente independentes, então o seu posto será igual a n, $p(I) = n$, assim, $p(E) = n$.

Matriz Aumentada ou Ampliada

Dado o sistema (SL_m), $AX = B$, a matriz aumentada ou ampliada do sistema, denotada por A_u, é dada ao acrescentarmos na matriz de coeficientes, A, a matriz dos termos independentes, B.

Notação: $A_u = [A \vdots B]$.

A_u é uma matriz de ordem $m \times (n+1)$, onde as n primeiras colunas são dadas pela matriz A e a última coluna é a matriz coluna dada por B. Por exemplo, a matriz aumentada do sistema

$$\begin{aligned} x + 3y + 5z &= 1 \\ x + 2y + 7z &= 2 \\ 2y + z &= 3 \end{aligned}$$

é dada por:

$$A_u = \begin{bmatrix} 1 & 3 & 5 \vdots 1 \\ 1 & 2 & 7 \vdots 2 \\ 0 & 2 & 1 \vdots 3 \end{bmatrix}.$$

Ao fazermos operações elementares na matriz A_u, estaremos mudando simultaneamente os elementos de B, ou seja, obtemos uma matriz $\tilde{A}_u = [\tilde{A} \vdots \tilde{B}]$. Com isso, podemos dizer que o sistema inicial foi mudado para um sistema da forma:

$$\tilde{A}X = \tilde{B}.$$

Sistemas Equivalentes

Dois sistemas, $AX = B$ e $\tilde{A}X = \tilde{B}$, são ditos equivalentes se as matrizes aumentadas dos mesmos, $A_u = [A \vdots B]$ e $\tilde{A}_u = [\tilde{A} \vdots \tilde{B}]$, são matrizes equivalentes.

Exemplo 8. Encontre as matrizes aumentadas dos sistemas lineares (SL_{g1}) e (SL_{g2}) dados abaixo e verifique se eles são equivalentes.

$$\begin{aligned} x+3y+5z&=1 \\ x+2y+7z&=2 \\ 2y+z&=3 \end{aligned} \quad (SL_{g1}) \quad \text{e} \quad \begin{aligned} x+3y+5z&=1 \\ -y+2z&=1 \\ z&=1 \end{aligned} \quad (SL_{g2}).$$

Solução. As matrizes aumentadas dos sistemas (SL_{g1}) e (SL_{g2}) são, respectivamente, dados por:

$$A_{u1} = \begin{bmatrix} 1 & 3 & 5 & \vdots 1 \\ 1 & 2 & 7 & \vdots 2 \\ 0 & 2 & 1 & \vdots 3 \end{bmatrix} \quad \text{e} \quad A_{u2} = \begin{bmatrix} 1 & 3 & 5 & \vdots 1 \\ 0 & -1 & 2 & \vdots 1 \\ 0 & 0 & 1 & \vdots 1 \end{bmatrix}.$$

Veja que, fazendo as três operações elementares, $l_2 \leftarrow -l_1 + l_2$, $l_3 \leftarrow l_3 + 2l_2$ e $l_3 \leftarrow \frac{l_3}{5}$ sobre a matriz A_{u1}, obtemos facilmente a matriz A_{u2}, isto é, $A_{u1} \xrightarrow[\text{elementares}]{\text{operações}} A_{u2}$. Assim, A_{u1} e A_{u2} são matrizes equivalentes.

Propriedade. Sistemas equivalentes possuem a mesma solução.

Assim, se a forma reduzida da matriz aumentada é muito parecida com a forma triangular ou diagonal, a propriedade sugere que ao resolver um sistema linear é suficiente resolver o sistema equivalente obtido após o escalonamento.

Dessa forma, usando métodos convencionais nos sistemas dados no **Exemplo 8**, verificamos que a solução, em ambos os casos, é $x = -7$, $y = 1$, $z = 1$. Isto é, $X = \begin{bmatrix} -7 \\ 1 \\ 1 \end{bmatrix}$, agora **pedimos que você verifique o resultado!**

Com o conteúdo teórico dado nas seções já apresentadas, estamos prontos para apresentar o método de solução usando escalonamento. Para isso, precisamos caracterizar os sistemas lineares!

2.2.3 Caracterização dos Sistemas Lineares

Um sistema linear pode ser:

a) **Possível** (compatível, consistente), se possui solução.

- *Determinado:* quando a solução é única;
- *Indeterminado:* quando há infinitas soluções.

b) **Impossível** (incompatível, inconsistente), se não possui solução.

Seja o sistema linear de m equações com n incógnitas da forma:

$$AX = B.$$

Usando o conceito de posto no caso das matrizes A e A_u $(p(A)$ e $p(A_u))$ caracterizamos as soluções do sistema linear da seguinte forma:

a) O sistema é possível se $p(A) = p(A_u)$.
Se $p = p(A)$, então o sistema é:

- Determinado, se $p = n$.
- Indeterminado, se $p < n$.

b) O sistema é impossível se $p(A) \neq p(A_u)$.

Exemplo 9. Caracterize os sistemas abaixo.

a) $\begin{aligned} -x + y &= 1 \\ -x + y &= 2 \end{aligned}$

b) $\begin{aligned} -x + y &= 1 \\ x + y &= -1 \end{aligned}$

Solução.

a) Considerando $m = 2$ e $n = 2$, o sistema pode ser colocado na seguinte forma matricial:

$$\overset{A}{\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}} \cdot \overset{X}{\begin{bmatrix} x \\ y \end{bmatrix}} = \overset{B}{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}$$

$$A_u = [A \vdots B] = \begin{bmatrix} -1 & 1 \vdots 1 \\ -1 & 1 \vdots 2 \end{bmatrix} \xrightarrow{l_2 \leftarrow -l_2 + l_1} \begin{bmatrix} -1 & 1 \vdots 1 \\ 0 & 0 \vdots -1 \end{bmatrix} = \tilde{A}_u$$

Pela propriedade do posto de matrizes equivalentes, temos que $p(A) = 1$ e $p(A_u) = 2$. Como $p(A_u) \neq p(A)$, então, o sistema é impossível.

b) Com $m = 2$ e $n = 2$, o sistema $AX = B$ é da forma:

$$\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$A_u = [A \vdots B] = \begin{bmatrix} -1 & 1\vdots & 1 \\ 1 & 1\vdots & -1 \end{bmatrix} \xrightarrow{op.elem.} \begin{bmatrix} -1 & 1\vdots & 1 \\ 0 & 2\vdots & 0 \end{bmatrix} = \tilde{A}_u .$$

Assim, $p(A) = 2$ e $p(A_u) = 2$, por serem iguais $(p = 2)$, o sistema é possível. Como $n = 2 = p$ teremos, então, um sistema possível e determinado.

O sistema equivalente ao sistema dado é:

$$\begin{matrix} -x + y = 1 \\ 2y = 0 \end{matrix} .$$

Assim,

$$\begin{matrix} x = -1 \\ y = 0 \end{matrix} .$$

Portanto,

$$X = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

é a solução do sistema dado.

Agora é sua vez!

Exercício 5. Com auxílio do conteúdo fornecido na **Seção 2.2**, interprete geometricamente os sistemas (a) e (b) dados no **Exemplo 9**.

Graus de Liberdade de um Sistema Linear

Quando um sistema linear é caracterizado como indeterminado, usa-se o conceito de graus de liberdade. O mesmo será denotado pela letra g.

Assim, se um sistema de m equações com n incógnitas da forma $AX = B$ é indeterminado, e se o posto da matriz de coeficientes é p, então $p(A) = p(A_u) = p < n$, isto é, $n - p > 0$.

Dessa forma, definimos $g = n - p$.

Observações:

- Os graus de liberdade do sistema, g, são sempre um valor positivo.
- O valor de g será associado ao número de variáveis livres do sistema.

Exemplo 10. Dado o sistema

$$-x + y = 1,$$

encontre os graus de liberdade, indique as variáveis livres e encontre a solução.

Solução. A representação matricial do sistema é

$$\begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = [1].$$

Nesse caso, $m = 1$, $n = 2$ e $p(A) = p(A_u) = 1$. **Você saberia dizer por quê?**

O sistema é possível e indeterminado; assim, $g = 2 - 1 = 1$. Isto é, temos uma variável livre. Essa variável pode ser escolhida como sendo qualquer uma das incógnitas do sistema x ou y. Sem perda de generalidade, escolheremos a segunda, y.

Quais são as soluções do sistema? Como o sistema possui infinitas soluções, consideramos $y = r \in \mathbb{R}$. Encontraremos todas as soluções por substituição direta:

$$\begin{aligned} x &= -1 + y = -1 + r \\ y &= r \end{aligned}$$

Representando matricialmente:

$$X = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 + r \\ r \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} r, \quad r \in \mathbb{R}.$$

Assim, o conjunto de soluções do sistema dado é:

$$\left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \,/ \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} r, \quad r \in \mathbb{R} \right\}.$$

Observação. Fazendo $r = 0$ ou $r = 2$, no sistema dado no **Exemplo 10**, teremos, respectivamente:

$$X_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \text{ ou } X_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Assim, X_1 e X_2 são duas soluções particulares do sistema dado.

Exercício Resolvido

1) Dado o sistema:

$$\begin{aligned} x_1 + x_2 - x_4 &= 0 \\ x_1 - x_3 + x_4 &= 2 \\ x_2 + x_3 - x_4 &= -3 \\ x_1 + x_2 - 2x_4 &= 1 \\ 2x_1 + x_2 - x_3 - x_4 &= 3 \end{aligned}$$

Encontre:

a) A forma matricial do sistema;

b) O posto e a nulidade da matriz de coeficientes;

c) Os graus de liberdade do sistema;

d) A caracterização do sistema;

e) Calcular a solução (se existir).

Solução.

a) Considerando $m = 5$ e $n = 4$, o sistema na forma $AX = B$ é dado por:

$$\begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & 0 & -2 \\ 2 & 1 & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ -3 \\ 1 \\ 3 \end{bmatrix}$$

b) Escalonando $A_u = [A \vdots B]$,

$$\begin{bmatrix} 1 & 1 & 0 & -1 & \vdots & 0 \\ 1 & 0 & -1 & 1 & \vdots & 2 \\ 0 & 1 & 1 & -1 & \vdots & -3 \\ 1 & 1 & 0 & -2 & \vdots & 1 \\ 2 & 1 & -1 & -1 & \vdots & 3 \end{bmatrix} \xrightarrow[l_5 \leftarrow -2l_1 + l_5]{\substack{l_2 \leftarrow -l_1 + l_2 \\ l_4 \leftarrow -l_1 + l_4}} \begin{bmatrix} 1 & 1 & 0 & -1 & \vdots & 0 \\ 0 & -1 & -1 & 2 & \vdots & 2 \\ 0 & 1 & 1 & -1 & \vdots & -3 \\ 0 & 0 & 0 & -1 & \vdots & 1 \\ 0 & -1 & -1 & 1 & \vdots & 3 \end{bmatrix} \xrightarrow[l_5 \leftarrow -l_2 + l_5]{l_3 \leftarrow l_2 + l_3}$$

$$\begin{bmatrix} 1 & 1 & 0 & -1 & \vdots & 0 \\ 0 & -1 & -1 & 2 & \vdots & 2 \\ 0 & 0 & 0 & 1 & \vdots & -1 \\ 0 & 0 & 0 & -1 & \vdots & 1 \\ 0 & 0 & 0 & -1 & \vdots & 1 \end{bmatrix} \xrightarrow[\substack{l_5 \leftarrow -l_4 + l_5 \\ -l_2}]{l_4 \leftarrow l_3 + l_4} \begin{bmatrix} 1 & 1 & 0 & -1 & \vdots & 0 \\ 0 & 1 & 1 & -2 & \vdots & -2 \\ 0 & 0 & 0 & 1 & \vdots & -1 \\ 0 & 0 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 0 & 0 & \vdots & 0 \end{bmatrix},$$

assim:

$$p = p(A) = 3 \text{ e } nul(A) = 1;$$

c) $g(A) = n - p = 4 - 3 = 1;$

d) Como $p(A) = p(A_u) = 3 < 4$, então o sistema é indeterminado e possui 1 variável livre;

e) O sistema equivalente é

$$\begin{aligned} x_1 + x_2 - x_4 &= 0 \\ x_2 + x_3 - 2x_4 &= -2 \\ x_4 &= -1 \end{aligned}$$

e, considerando x_3 como sendo a variável livre, isto é, $x_3 = r \in \mathbb{R}$, temos:

$$\begin{aligned} x_4 &= -1 \\ x_3 &= r \end{aligned} \quad \Rightarrow \quad \begin{aligned} x_2 &= 2 - r + 2(-1) = -4 - r \\ x_1 &= -x_2 + x_4 = 3 + r \end{aligned}.$$

Assim, a solução X é igual a:

$$X = \begin{bmatrix} 3 + r \\ -4 - r \\ r \\ -1 \end{bmatrix}, \quad r \in \mathbb{R},$$

e o conjunto solução, S, pode ser colocado na seguinte forma:

$$S=\left\{X=\begin{bmatrix}3+r\\-4-r\\r\\-1\end{bmatrix}, r\in\mathbb{R}\right\}=\left\{\begin{bmatrix}3\\-4\\0\\-1\end{bmatrix}+\begin{bmatrix}1\\-1\\1\\0\end{bmatrix}r, r\in\mathbb{R}\right\}.$$

A solução do exercício nos fornece o método comumente conhecido como o **método de Gauss** .

2.2.4 Método de Gauss para Solução de Sistemas Lineares

Ao resolvermos os sistemas no **Exemplo 10** e o exercício resolvido (1), optamos por escolher uma variável livre (ou várias, se fosse o caso) e, por substituição a partir dela (ou delas), encontramos as outras variáveis, usando o sistema equivalente. Esse método é conhecido como o método de Gauss para sistemas lineares.

Observemos com detalhes o que foi feito no exercício resolvido (1). Após o escalonamento, chegamos ao seguinte sistema equivalente, onde a quarta incógnita, x_4, ficou dada em forma explícita:

$$x_1+x_2-x_4=0 \quad (1)$$
$$x_2+x_3-2x_4=-2 \quad (2)$$
$$x_4=-1 \quad (3)$$

Com a equação (3), resolvemos x_1 e x_2, e considerando a variável livre:

$$x_3=r\in\mathbb{R} \quad (4)$$

Assim, substituindo (3) e (4) em (1) e (2), conseguimos o seguinte resultado:

$$x_1+x_2-(-1)=0$$
$$x_1+x_2=-1$$

$$x_2+r-2(-1)=-2 \quad (5)$$
$$x_2+r=-4$$
$$x_2=-4-r$$

$$x_1+x_2=-1 \quad (6)$$
$$x_1-4-r=-1$$
$$x_1=3+r$$

Considerando que $X = [x_1 \quad x_2 \quad x_3 \quad x_4]^t$ e os resultados dados em (6), (5), (4) e (3), temos a solução geral:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3+r \\ -4-r \\ r \\ -1 \end{bmatrix}, \quad r \in \mathbb{R}.$$

Agora é sua vez!

Exercício 6. Duas soluções do sistema no exercício resolvido são $X_1 = \begin{bmatrix} 4 \\ -5 \\ 1 \\ -1 \end{bmatrix}$ e $X_2 = \begin{bmatrix} 5 \\ -6 \\ 2 \\ -1 \end{bmatrix}$. Para quais valores de r isso acontece?

Observação. É possível escolher qualquer uma das incógnitas como variável livre.

No exercício resolvido, nada impede de escolhermos x_1 ou x_2 como variáveis livres. Vejamos, como exemplo:

Se a variável livre escolhida for $x_1 = s \in \mathbb{R}$, teremos

em (1)

$$\begin{aligned} &x_1 + x_2 - x_4 = 0 \\ &s + x_2 - (-1) = 0 \\ &x_2 = -1 - s \end{aligned}$$

em (2)

$$\begin{aligned} &x_2 + x_3 - 2x_4 = -2 \\ &-s - 1 + x_3 - 2(-1) = -2 \\ &x_3 = s + 1 - 4 \\ &x_3 = -3 - s \end{aligned}$$

Assim, obtemos a seguinte solução geral:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} s \\ -1-s \\ -3+s \\ -1 \end{bmatrix}, \quad s \in \mathbb{R}.$$

O conjunto solução, nesse caso, é:

$$S = \left\{ \begin{bmatrix} 0 \\ -1 \\ -3 \\ -1 \end{bmatrix} + s \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \quad s \in \mathbb{R} \right\}.$$

Considerando $s = 4$, obtemos a solução particular $X_1 = \begin{bmatrix} 4 \\ -5 \\ 1 \\ -1 \end{bmatrix}$, observe que é a mesma solução encontrada anteriormente, em X_1, do Exercício 6.

Agora é sua vez!

Exercício 7. Qual é o valor de s para encontrarmos a outra solução, X_2, do Exercício 6?

Soluções como Combinação Linear de Vetores

Agora, suponha um sistema indeterminado $AX = B$, onde $n = 4$ e $p = 2$, que tenha duas variáveis livres ($g = 2$) e cuja solução geral fique explicitada como sendo, por exemplo:

$$X = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix} + r \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ -2 \\ 0 \end{bmatrix}, \quad r, s \in \mathbb{R}.$$

A solução desse exemplo hipotético é colocada como a soma de uma solução particular não nula e uma combinação linear de dois vetores. Isso indica que, em geral, podemos expressar a solução de um sistema linear indeterminado qualquer como uma soma de uma solução particular e uma combinação linear de vetores (essa combinação terá terá um número de vetores igual ao número de variáveis livres). Ou seja, toda vez que o sistema for indeterminado, a solução geral pode ser caracterizada por:

$$X = X_p + t_1 X_1 + t_2 X_2 + \ldots + t_g X_g,$$

em que as soluções X_1, X_2, ..., X_g serão denominadas **soluções básicas** e X_p é a **solução particular.**

Exemplificaremos esse resultado nos dois exercícios seguintes, sendo que o primeiro será resolvido detalhadamente.

Exercícios Resolvidos

2) Verifique que o seguinte sistema é indeterminado e que a solução pode ser colocada como a soma de uma solução particular e uma combinação linear de soluções básicas:

$$\begin{aligned}
&10x_2 - 4x_3 + x_4 = 1\\
&x_1 + 4x_2 - x_3 + x_4 = 2\\
&3x_1 + 2x_2 + x_3 + 2x_4 = 5\\
&-2x_1 - 8x_2 + 2x_3 - 2x_4 = -4\\
&x_1 - 6x_2 + 3x_3 = 1
\end{aligned}$$

Solução. Considerando $m = 5$ e $n = 4$, expressamos o sistema na seguinte forma matricial:

$$\begin{bmatrix} 0 & 10 & -4 & 1 \\ 1 & 4 & -1 & 1 \\ 3 & 2 & 1 & 2 \\ -2 & -8 & 2 & -2 \\ 1 & -6 & 3 & 0 \end{bmatrix} X = \begin{bmatrix} 1 \\ 2 \\ 5 \\ -4 \\ 1 \end{bmatrix}.$$

Usando o método de Gauss,

$$A_u = \begin{bmatrix} 0 & 10 & -4 & 1 & \vdots & 1 \\ 1 & 4 & -1 & 1 & \vdots & 2 \\ 3 & 2 & 1 & 2 & \vdots & 5 \\ -2 & -8 & 2 & -2 & \vdots & -4 \\ 1 & -6 & 3 & 0 & \vdots & 1 \end{bmatrix} \xrightarrow{l_1 \leftrightarrow l_2} \begin{bmatrix} 1 & 4 & -1 & 1 & \vdots & 2 \\ 0 & 10 & -4 & 1 & \vdots & 1 \\ 3 & 2 & 1 & 2 & \vdots & 5 \\ -2 & -8 & 2 & -2 & \vdots & -4 \\ 1 & -6 & 3 & 0 & \vdots & 1 \end{bmatrix}$$

$$\xrightarrow[\substack{l_4 \leftarrow l_4 + 2l_1 \\ l_5 \leftarrow l_5 - l_1}]{l_3 \leftarrow l_3 - 3l_1} \begin{bmatrix} 1 & 4 & -1 & 1 & \vdots & 2 \\ 0 & 10 & -4 & 1 & \vdots & 1 \\ 0 & -10 & 4 & -1 & \vdots & -1 \\ 0 & 0 & 0 & 0 & \vdots & 0 \\ 0 & -10 & 4 & -1 & \vdots & -1 \end{bmatrix} \xrightarrow[\substack{l_5 \leftarrow l_2 + l_5 \\ \frac{l_2}{10}}]{l_3 \leftarrow l_2 + l_3} \begin{bmatrix} 1 & 4 & -1 & 1 & \vdots & 2 \\ 0 & 1 & \frac{-2}{5} & \frac{1}{10} & \vdots & \frac{1}{10} \\ 0 & 0 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 0 & 0 & \vdots & 0 \end{bmatrix} = \tilde{A}_u$$

Como $\mathrm{p} = \mathrm{p(A)} = \mathrm{p(A}_u) = 2 < 4$, então o sistema é indeterminado.

Os graus de liberdade do sistema são $\mathrm{g} = 2$. Assim sendo, haverá duas variáveis livres.

O sistema equivalente é dado por:

$$x_1 + 4x_2 - x_3 + x_4 = 2 \qquad (1)$$

$$x_2 - \frac{2}{5}x_3 + \frac{1}{10}x_4 = \frac{1}{10} \qquad (2)$$

Sejam x_3 e x_4 as variáveis livres, isto é,

$$x_3 = s\text{, } x_4 = r \text{ com } s, r \in \mathbb{R}. \qquad (3)$$

Substituindo (3) em (2), obtemos:

$$x_2 - \frac{2}{5}x^3 + \frac{1}{10}x_4 = \frac{1}{10}$$

$$x_2 = \frac{1}{10} + \frac{2}{5}s - \frac{1}{10}r \qquad (4)$$

Similarmente, substituindo (3) e (4) em (1):

$$x_1 + 4x_2 - x_3 + x_4 = 2$$

$$x_1 = 2 - 4\left(\frac{1}{10} + \frac{2}{5}s - \frac{1}{10}r\right) + s - r \qquad (5)$$

$$x_1 = \frac{8}{5} - \frac{3}{5}s - \frac{3}{5}r$$

A solução geral é:

$$\mathrm{X} = \begin{bmatrix} \frac{8}{5} \\ \frac{1}{10} \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} \frac{-3}{5} \\ \frac{2}{5} \\ 1 \\ 0 \end{bmatrix} + r\begin{bmatrix} \frac{-3}{5} \\ \frac{-1}{10} \\ 0 \\ 1 \end{bmatrix}, \quad s, r \in \mathbb{R}$$

Dessa forma, a solução geral é apresentada como a soma de uma solução particular,

$\begin{bmatrix} \frac{8}{5} \\ \frac{1}{10} \\ 0 \\ 0 \end{bmatrix}$, e uma combinação linear dos vetores (matrizes) coluna $\begin{bmatrix} \frac{-3}{5} \\ \frac{2}{5} \\ 1 \\ 0 \end{bmatrix}$ e $\begin{bmatrix} \frac{-3}{5} \\ \frac{-1}{10} \\ 0 \\ 1 \end{bmatrix}$.

3) Verifique se o sistema dado a seguir é indeterminado e encontre a solução geral.

$$\begin{aligned} x_1 + 4x_2 - 6x_3 &= 1 \\ -2x_1 - 8x_2 + 12x_3 &= -2 \\ 2x_1 - 3x_2 + x_3 &= 0 \end{aligned}$$

Solução. O sistema possui $m = 3$ equações e $n = 3$ incógnitas e a representação matricial dele é dada por:

$$\begin{bmatrix} 1 & 4 & -6 \\ -2 & -8 & 12 \\ 2 & -3 & 1 \end{bmatrix} X = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$$

Após escalonamento (usando o método de Gauss), obtemos que $g = 1$, e considerando x_3 como variável livre, encontramos a seguinte solução geral:

$$X = \begin{bmatrix} \frac{3}{11} \\ \frac{2}{11} \\ 0 \end{bmatrix} + t \begin{bmatrix} \frac{14}{11} \\ \frac{13}{11} \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

2.2.5 Sistemas Homogêneos

Quando $B = O$ no sistema dado por $AX = B$, o sistema é dito ser homogêneo. Isto é, um sistema homogêneo é um sistema linear da forma:

$$AX = O, \quad (SL_h)$$

onde SL_h indica ser um Sistema Linear homogêneo.

Observe que, nesse caso, o posto da matriz é igual ao posto da matriz aumentada, uma vez que a última coluna da matriz aumentada é nula e permanecerá inalterada após as operações elementares. Dessa forma, o sistema é sempre possível!

Consideremos que A possua n linhas, então, se o sistema for determinado ($\text{p} = n$), haverá n linhas linearmente independentes e obteremos a solução nula, conhecida como **solução trivial**. Você pode verificar, ao fazer substituição direta, após o escalonamento, que $\text{X} = \text{O}$. Já no caso do sistema ser indeterminado, haverá pelo menos uma solução básica diferente da trivial a partir da qual (ou das quais) todas as outras soluções são geradas.

Poderia justificar por quê?

Observação: Observe que, pelo desenvolvido na Seção anterior, a solução particular nesse caso será sempre nula.

Vejamos os seguintes exemplos.

Exemplo 11. Encontre a solução do seguinte sistema homogêneo:

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 0 \\ x_1 + 3x_2 &= 0 \\ x_1 - x_2 - x_3 &= 0 \end{aligned}$$

Solução. Após escalonamento, temos o seguinte sistema equivalente:

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 0 \\ x_2 + x_3 &= 0 \\ x_3 &= 0 \end{aligned}$$

Sendo que $n = 3$ e $\text{p} = 3$, então, o sistema é determinado. Assim, a solução é a solução trivial. Confirmamos isso resolvendo por substituição direta a partir da última equação, conseguindo:

$$\text{X} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Você pode facilmente verificar esse resultado.

Exemplo 12. Encontre as soluções básicas do sistema homogêneo dado por:

$$\begin{bmatrix} 1 & 4 & -6 \\ -2 & -8 & 12 \\ 2 & -3 & 1 \end{bmatrix} X = O.$$

Solução. Você pode observar que o sistema dado corresponde à parte homogênea do sistema dado no último exercício resolvido. Nesse exercício, obtivemos um sistema indeterminado com um grau de liberdade, isto é, o sistema homogêneo possui uma solução básica da forma:

$$X_1 = \begin{bmatrix} \frac{14}{11} \\ \frac{13}{11} \\ 1 \end{bmatrix}.$$

Podemos expressar a solução geral na forma:

$$X = tX_1 = t\begin{bmatrix} \frac{14}{11} \\ \frac{13}{11} \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Propriedades

Dado o sistema linear $AX = B$ com parte homogênea $AX = O$, listamos a seguir as seguintes propriedades.

Propriedade 1. A solução de um sistema linear é a soma da solução da parte homogênea, X_h, com uma solução particular, X_p, do sistema não homogêneo.

Prova. Por hipótese, X_h é solução do sistema homogêneo, isto é, $AX_h = O$ e X_p é uma solução particular do sistema não homogêneo, ou seja, $AX_p = B$.

Vejamos se $X_p + X_h$ é solução do sistema linear:

$$A(X_p + X_h) = AX_p + AX_h.$$

Sendo que

$$AX_h = O \text{ e } AX_p = B,$$

então,

$$A(X_p + X_h) = O + B$$
$$= B$$

Daí que $X_p + X_h$ é solução do sistema linear.

Exemplo 13. Agora vamos verificar a propriedade no sistema a seguir:

$$x_1 + 2x_2 - x_3 = 1$$
$$x_1 + 3x_2 = 0$$
$$x_1 - x_2 - x_3 = 0$$

Solução. A forma matricial do sistema dado é:

$$\begin{bmatrix} 1 & 2 & -1 \\ 1 & 3 & 0 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad (1)$$

O respectivo sistema homogêneo é

$$AX = O, \qquad (2)$$

cujo sistema foi fornecido no **Exemplo 11.**

Ao resolvermos (2), obtemos:

$$x_1 = x_2 = x_3 = 0 \Rightarrow X_h = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Ao resolvermos (1), temos

$$X_p = \begin{bmatrix} -1 \\ 1/3 \\ -4/3 \end{bmatrix}.$$

Assim,

$$X = X_p + X_1 = \begin{bmatrix} -1 \\ 1/3 \\ -4/3 \end{bmatrix}.$$

e a propriedade fica satisfeita.

Agora é sua vez!

Exercício 8. Anteriormente, resolvemos o sistema $AX = B$ da forma:

$$\begin{bmatrix} 0 & 10 & -4 & 1 \\ 1 & 4 & -1 & 1 \\ 3 & 2 & 1 & 2 \\ -2 & -8 & 2 & -2 \\ 1 & -6 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \\ -4 \\ 1 \end{bmatrix} \qquad (1)$$

Com solução geral:

$$X = \begin{bmatrix} \frac{8}{5} \\ \frac{1}{10} \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} \frac{-3}{5} \\ \frac{2}{5} \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} \frac{-3}{5} \\ \frac{-1}{10} \\ 0 \\ 1 \end{bmatrix}, \quad s, t \in \mathbb{R}.$$

A solução do sistema homogêneo respectivo, $AX = O$, é da forma:

$$X_h = s\begin{bmatrix} \frac{-3}{5} \\ \frac{2}{5} \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} \frac{-3}{5} \\ \frac{-1}{10} \\ 0 \\ 1 \end{bmatrix}, \quad s, t \in \mathbb{R}. \qquad (2)$$

a) Obtenha:

i) X_n para $t = s = 1$.

ii) X_h para $t = 2$ e $s = -1$.

b) Verifique a **Propriedade 1** usando as soluções obtidas nos itens (i) e (ii). *Utilize o mesmo conjunto de parâmetros em cada caso.*

Propriedade 2. Dado um sistema homogêneo com solução diferente da trivial:

a) Se X e Y são soluções, então $X + Y$ também é solução.

b) Se X é solução, então αX também é solução, para α qualquer escalar real.

Prova. Como exercício, demonstre a **Propriedade 2.**

Observação. A Propriedade 2 também pode ser formulada da seguinte forma:

Se X e Y são soluções do sistema $AX = O$, então, $\alpha X + \beta Y$ também é uma solução do sistema.

2.3 Decomposição LU

Algumas vezes, é muito útil fatorar um número natural em um produto de outros números naturais (por exemplo, $6 = 2 \cdot 3$, $45 = 3 \cdot 3 \cdot 5$ etc.). Dessa mesma forma, pode ser muito útil fatorar ou decompor uma matriz como o produto de duas ou mais matrizes. Um exemplo disso é o seguinte:

$$\begin{bmatrix} 3 & -1 \\ 9 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 0 & -2 \end{bmatrix}.$$

O que você observa? Há alguma característica especial em cada uma das matrizes da fatoração dada?

A ***decomposição de matrizes*** é geralmente usada na resolução de sistemas de equações lineares pelo método da eliminação de Gauss, e particularmente adequada para **implementação em computadores**.

Embora o escopo desse texto aborde apenas a decomposição LU, você deve saber que existem outros tipos de fatoração de matrizes, igualmente úteis: fatoração QR, fatoração de Cholesky etc. Esses tópicos são muito ricos e existem várias referências que apresentam esses assuntos.

Vejamos como podemos realizar tal fatoração, seja qual for a ordem da matriz. Para tal, considere um sistema de equações lineares da forma $AX = B$, em que A é uma matriz de ordem n.

Exemplo 14. Seja A uma matriz de ordem $n = 3$ da forma:

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix},$$

escalone-a até obter uma matriz triangular superior.

Solução. O escalonamento pedido pode seguir o seguinte processo:

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} = U.$$

Lembra-se, no capítulo anterior, do desenvolvimento realizado para calcularmos a matriz inversa? Pois bem, desta vez, um procedimento similar servirá para decompormos uma matriz qualquer.

Seguindo com o exemplo dado acima, podemos observar que no processo de decomposição podem ser associadas três matrizes elementares E_1, E_2, E_3 tais que

$$E_3E_2E_1A = U.$$

Você pode verificar que essas matrizes elementares são dadas por:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}.$$

Agora é sua vez!

Exercício 9. a) Verifique se o produto $E_3E_2E_1A = U$.

b) Considere $L^{-1} = E_3E_2E_1$, verifique matricialmente que $A = LU$, também encontre L - a mesma será uma matriz triangular inferior com elementos diagonais iguais a 1.

Você observará que o método está sendo introduzido numa forma muito prática, ao estudar o **Exemplo 14** e resolver o Exercício 9. Pois bem, essa prática fornecerá a metodologia para calcular a decomposição de uma matriz quadrada qualquer.

Suponha que A é uma matriz de ordem n que permite ser escalonada sem precisar da permutação das suas linhas:

$$A \xrightarrow[\text{elementares}]{\text{operações}} U,$$

isto é, existem matrizes elementares $E_1, E_2, \ldots, E_k$ tais que:

$$E_k \ldots E_2 E_1 A = U.$$

Como visto anteriormente, cada matriz E_i é não singular; assim, o produto $E_k \ldots E_2 E_1$ é também não singular, isto é, existe $(E_k \ldots E_2 E_1)^{-1}$, tal que

$$A = (E_k \ldots E_2 E_1)^{-1} U.$$

A matriz $L = (E_k \ldots E_2 E_1)^{-1}$ é uma matriz triangular inferior com elementos diagonais iguais a um (poderia dizer por quê?).

Dessa forma, obtemos a decomposição da matriz A como o produto de uma matriz triangular inferior L por uma matriz triangular superior U, isto é,

$$A = LU.$$

Mais detalhes da metodologia, assim como um código computacional da mesma, podem ser encontrados no ambiente de aprendizagem da nossa disciplina.

É importante ressaltar que o procedimento mostrado anteriormente foi desenvolvido sem a permutação das linhas da matriz. Em geral, ao escalonarmos uma matriz, poderemos precisar desse tipo de operação elementar. Nesse caso, você pode recorrer à leitura das referências dadas pelos livros de Leon ou Poole.

A decomposição LU serve para calcular a matriz inversa, o determinante de uma matriz e para resolver sistemas lineares. É neste último ponto, a resolução de sistemas lineares, que enfocaremos a praticidade desse método no conteúdo desta Seção.

Dado o sistema linear

$$AX = B,$$

e usando a decomposição $A = LU$, temos:

$$LUX = B.$$

Encontraremos X ao inserirmos uma nova variável vetorial Y que resolva o sistema

$$LY = B.$$

Assim, ao termos encontrado Y, resolvemos:

$$Y = UX \quad \text{ou} \quad UX = Y.$$

Nas duas situações, você pode reparar que foram resolvidos dois sistemas numa forma recursiva. Vejamos isso com o seguinte exemplo.

Exemplo 15. Resolver o sistema $AX = B$ usando decomposição LU de matrizes. Use a matriz $A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix}$ e $B = \begin{bmatrix} 2 \\ 0 \\ -5 \end{bmatrix}$.

Solução. Com o método acima descrito e sabendo que

$$A = LU,$$

isto é :

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix},$$

temos $LY = B$ ou

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} Y = \begin{bmatrix} 2 \\ 0 \\ -5 \end{bmatrix}, \text{ onde } Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

Assim,

$$\begin{aligned} y_1 &= 2 \\ y_2 &= -4 \\ y_3 &= -5 + y_1 + 2y_2 = -11 \end{aligned} \qquad \Rightarrow Y = \begin{bmatrix} 2 \\ -4 \\ -11 \end{bmatrix}$$

Agora, resolvendo $UX = Y$,

$$\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} X = \begin{bmatrix} 2 \\ -4 \\ -11 \end{bmatrix}$$

$$2x_3 = -11 \quad \Rightarrow \quad x_3 = \frac{-11}{2}$$

$$-3x_2 = -4 + 3x_3 \quad \Rightarrow \quad x_2 = \frac{41}{6} \quad \Rightarrow \quad X = \begin{bmatrix} \frac{35}{6} \\ \frac{41}{6} \\ \frac{-11}{2} \end{bmatrix}$$

$$2x_1 = 2 - x_2 - 3x_3 = \frac{70}{6} \quad \Rightarrow \quad x_1 = \frac{35}{6}$$

A solução do sistema é dada pelo vetor $X = \begin{bmatrix} \frac{35}{6} \\ \frac{41}{6} \\ \frac{-11}{2} \end{bmatrix}$.

Pratique agora!

Exercício 10. Verifique a resposta do **Exemplo 15** usando outro método.

Observações:

1) A decomposição LU é muito útil para resolver sistemas de ordem $n \geq 4$, em que será necessário fazer muitos cálculos.

2) Uma forma de reduzir os cálculos à mão é usar um Sistema Algébrico Computacional, conhecido na literatura como CAS (Computational Algebric System). Os softwares livres mais comuns são Scilab e Octave. Dentre os comerciais, podemos citar o Matlab (ou Matrix Laboratory), Maple e Mathematica.

Exercícios Propostos

1) Para cada sistema dado, encontre:

   i) a forma matricial do sistema;

   ii) o posto e a nulidade da matriz de coeficientes;

iii) os graus de liberdade do sistema;

iv) a caracterização do sistema;

v) a solução, se existir.

a)
$$\begin{aligned} 2x+y+3z+4w&=1\\ 3x+2y+4z+5w&=1\\ 2x+2y+z&=0\\ -2x+y+3z+w&=-1 \end{aligned}$$

b)
$$\begin{aligned} x_1+x_2+2x_3-x_4&=0\\ 2x_1+3x_2-2x_4&=0\\ -x_1+x_2+2x_3+2x_4&=0\\ x_1+2x_2+x_3+x_4&=0 \end{aligned}$$

c)
$$\begin{aligned} x+3z&=1\\ 2x-y+4z&=1\\ -6x-y-17z&=-1 \end{aligned}$$

2) Quais dos sistemas dados são possíveis ou impossíveis? Se for o caso, obtenha a solução geral e explicite as soluções básicas dos sistemas homogêneos relativos aos sistemas.

a)
$$\begin{aligned} 2x+y-2z+3w&=1\\ 3x+2y-z+2w&=4\\ 3x+3y+3z-3w&=5 \end{aligned}$$

b)
$$\begin{aligned} x_1-2x_2+3x_3-2x_4&=0\\ 3x_1-7x_2-2x_3+4x_4&=0\\ 4x_1+3x_2+5x_3+2x_4&=0 \end{aligned}$$

c)
$$\begin{aligned} x+2y-3z&=-1\\ 3x-y+2z&=7\\ 5x+3y-4z&=2 \end{aligned}$$

d)
$$\begin{aligned} x+2y-3z&=0\\ 2x+5y+2z&=0\\ 3x-y-4z&=0 \end{aligned}$$

3) Dado o sistema

$$\begin{aligned} x+y-z&=1\\ 2x+3y+az&=3\\ x+ay+3z&=2 \end{aligned}$$

Determine os valores de "a" para termos um sistema incompatível ou compatível (determinado ou indeterminado).

4) Se possível, faça a decomposição LU de cada uma das matrizes de coeficientes dos sistemas dados no Exercício 1. Resolva os sistemas usando o método de decomposição LU para sistemas lineares.

5) Aplique a decomposição LU para resolver os sistemas da forma $AX = B$ dados a seguir.

$$A = \begin{bmatrix} -2 & 1 \\ 2 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

$$A = \begin{bmatrix} 0 & 0 & 1 & 2 \\ -1 & 1 & 3 & 2 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}.$$

Resumo

Neste capítulo, trabalhamos conceitos e métodos fundamentais para resolver sistemas lineares de tipo retangular, isto é, aqueles em que a matriz de coeficientes tem um número de linhas não necessariamente igual ao número de colunas. Fornecemos os conceitos que podem ser aplicados em sistemas de ordem maior ou igual a quatro, onde os métodos convencionais ficam difíceis de serem aplicados.

Os conceitos foram trabalhados em paralelo com exemplos e exercícios resolvidos e propostos. Para entendimento do conteúdo, é necessária a prática da solução dos exercícios e a persistência nos detalhes da teoria fornecida nocapítulo anterior e neste próprio. Mais adiante, você poderá perceber quanto do aprendido será resgatado no capítulo que segue. Esperamos que você tenha entendido tudo quanto foi fornecido neste capítulo, empolgando-se na metodologia utilizada!

Bibliografia Comentada

BOLDRINI, José et. al. *Álgebra linear*. 3. ed. São Paulo: Harbra, 1980.

> Com esse livro, você pode pesquisar e aprofundar alguns dos conteúdos aqui fornecidos.

LEON, Steven J. *Álgebra linear com aplicações*. 4. ed. Rio de Janeiro: LTC, 1999.

POOLE, David. *Álgebra linear*. São Paulo: Pioneira Thomson Learning, 2004.

> Nessa referência, você encontrará exercícios da última Seção deste Capítulo e aplicações muito interessantes dos Capítulos aqui apresentados. Comentamos, por exemplo, o uso do código de barras, que é uma dessas aplicações em nosso cotidiano encontradas nesse livro. Fica, então, como convite, fazer uma leitura a respeito.

RORRES, Anton. *Álgebra linear com aplicações.* 8. ed. Porto Alegre: Bookman, 2001.

> Nesse livro, você pode encontrar uma gama de exercícios para acrescentar à sua prática.

SANTOS, Reginaldo. *Um curso de geometria analítica e álgebra linear.* Belo Horizonte: Ed. UFMG, 2001.

> Essa referência é muito útil para aplicar os conceitos usando alguns recursos computacionais.

Capítulo 3

Espaços Vetoriais

Capítulo 3
Espaços Vetoriais

A estrutura de espaço vetorial, que será tratada neste capítulo, é o conceito básico da Álgebra Linear. Um espaço vetorial não é outra coisa que um conjunto de vetores que satisfazem um conjunto de regras (axiomas). Este conceito proporciona unidade e precisão aos assuntos essenciais da Matemática. Após introduzir o sistema axiomático e as propriedades fundamentais, estaremos em condições de entender as idéias mais importantes que se derivam deste conceito.

3.1 Introdução

3.1.1 Definição

Em álgebra abstrata, um corpo é uma estrutura algébrica na qual estão definidas as operações de adição e multiplicação, que cumprem com as propriedades associativa, comutativa e distributiva.

Seja V um conjunto não vazio, K um corpo '+' e '*', duas operações que chamaremos soma e produto por escalar, respectivamente. Dizemos que o objeto $(V, \mathrm{K}, '+', '*')$ é um *espaço vetorial* (EV) (ou que V é um espaço vetorial sobre o corpo K) se, e somente se, verificam-se os seguintes axiomas:

EV1) A soma é uma composição interna em V, isto é, se $\mathbf{u}, \mathbf{v} \in V$, então $\mathbf{u} + \mathbf{v} \in V$.

EV2) A multiplicação de $u \in V$ por um escalar é uma composição externa em V, isto é, se α é um escalar em K e $\mathbf{u} \in V$, então $\alpha \mathbf{u} \in V$.

EV3) ...

$\vdots$

A esta lista segue uma série de propriedades que completam a definição de espaço vetorial, que omitiremos por enquanto (não se preocupem, elas serão relacionadas depois).

Isto é o que você pode esperar, em geral, quando se procura por uma definição de espaço vetorial nos textos de Álgebra Linear.

Para o leitor, que está começando com AL, esta definição pode ser formal demais para introduzir a idéia de EV. Em algum momento, deverá valorizar o rigor matemático da forma como a definição está sendo colocada acima, e gradativamente adotá-lo.

Agora, tomaremos um caminho alternativo para entender este conceito com maior facilidade. Valer-nos-emos, para isso, de exemplos de conjuntos (que são espaços vetoriais) com os quais você já tem alguma familiaridade.

Consideremos os seguintes conjuntos: o conjunto de vetores no plano $\mathbb{R}^2$, o conjunto formado pelas funções $f: \mathbb{R} \to \mathbb{R}$, denotado por $F(\mathbb{R})$, e o conjunto das matrizes quadradas de ordem m com coeficientes reais, que denotaremos por $M_{m \times m}$.

O matemático Hermann Grassmann (1809-1877) é geralmente creditado como o primeiro a introduzir a idéia de EV, em 1844. Porém, uma pessoa que realmente o estudou foi o matemático italiano Giuseppe Peano (1852-1932), que em seu livro Cálculo Geométrico, tornou claro o trabalho anterior de Grassmann e descreveu as propriedades para um EV da maneira como hoje o conhecemos. O livro de Peano é também digno de nota por introduzir operações em conjuntos, cujas anotações são: $\cup$, $\cap$ e $\in$ (que ainda usamos).

3.1.2 Vetores no Plano

Vetores não-nulos no plano podem ser representados geometricamente por segmentos orientados. Essa representação geométrica nos ajudaram a visualizar como as operações de multiplicação de um vetor por um escalar e de soma de vetores funcionam em $\mathbb{R}^2$. Consideremos um ponto (x_1, x_2) no plano, que não esteja situado na origem. É possível associar-lhe um segmento orientado com origem em $(0,0)$ e extremo em (x_1, x_2), que representamos como sendo o vetor $\mathbf{x} \in \mathbb{R}^2$, onde x_1 e x_2 são as respectivas componentes.

Vamos ao que nos interessa. Consideremos agora dois vetores $\mathbf{x}$ e $\mathbf{y}$, e um escalar $\alpha \in \mathbb{R}$. Sabemos que para cada vetor $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ e para cada escalar α, o produto $\alpha\mathbf{x}$ é definido por:

$$\alpha\mathbf{x} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \end{pmatrix},$$

que é um outro vetor que possui a mesma direção que $\mathbf{x}$ e que pode ter o sentido e comprimento mudado dependendo do valor de α.

Por outro lado, a soma de dois vetores $\mathbf{x}$ e $\mathbf{y}$ é definida como:

$$\mathbf{x}+\mathbf{y}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}+\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}=\begin{pmatrix} x_1+y_1 \\ x_2+y_2 \end{pmatrix}.$$

Esta operação pode ser resolvida graficamente, utilizando a regra do paralelogramo, então $\mathbf{x}+\mathbf{y}$ (o segmento orientado sobre a diagonal do paralelogramo) é representado por outro vetor, cuja origem coincide com a origem de $\mathbf{x}$ e extremo com o de $\mathbf{y}$.

Salientamos o fato de que $\alpha\mathbf{x} \in \mathbb{R}^2$ e $\mathbf{x}+\mathbf{y} \in \mathbb{R}^2$.

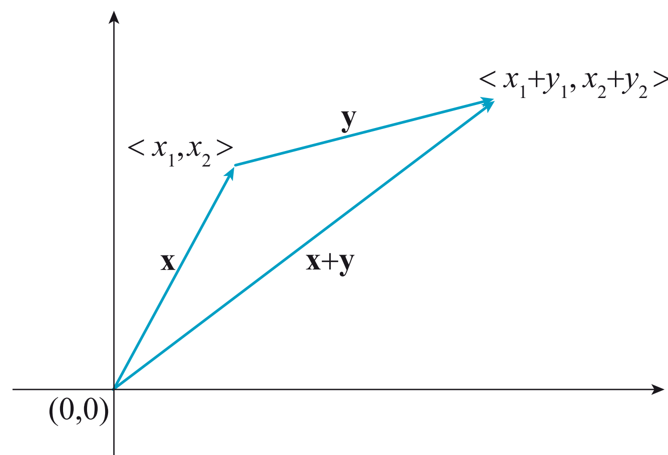


Figura 3.1 – A soma de $\mathbf{x}$ e $\mathbf{y}$ e o produto $\alpha\mathbf{x}$.

3.1.3 Funções Reais

A soma de duas funções f e g de $F(\mathbb{R})$ (funções com domínio nos reais) é definida como sendo a função $(f+g) \in F(\mathbb{R})$ dada por:

$$(f+g)(x)=f(x)+g(x).$$

Note também que, se $\alpha \in \mathbb{R}$, podemos multiplicar a função f pelo escalar α, da seguinte forma:

$$(\alpha f)(x)=\alpha(f(x)),$$

resultando, por sua vez, num elemento de $F(\mathbb{R})$. Mais uma vez, a soma e o produto por escalar originam funções de $F(\mathbb{R})$.

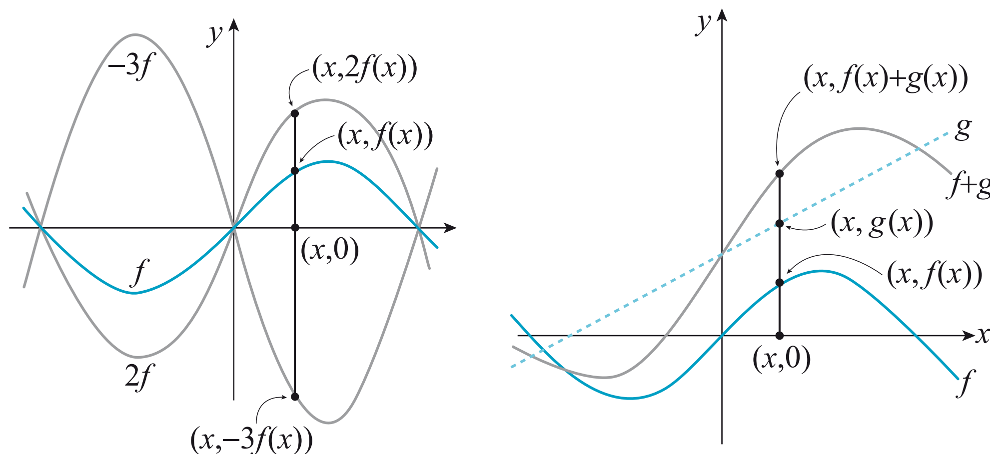


Figura 3.2 - A soma de funções e o produto por escalar.

3.1.4 Matrizes

Podemos somar duas matrizes quadradas de ordem n $(M_{n\times n})$ (as quais denotamos como $A = (a_{ij})_{n\times n}$ e $B = (b_{ij})_{n\times n}$), colocando $A + B = (a_{ij} + b_{ij})_{n\times n}$, que é também uma matriz de $M_{n\times n}$.

Com relação à multiplicação de $A = (a_{ij})_{n\times n}$ por um escalar $\alpha \in \mathbb{R}$, é natural definirmos $\alpha A = (\alpha a_{ij})_{n\times n}$, que também pertence a $M_{n\times n}$.

Você saberia dizer o que estes conjuntos citados anteriormente, com as operações de adição entre seus próprios elementos, por um lado, e a multiplicação de seus elementos por escalares, têm em comum?

Para começar, você pode ter percebido que: em todos os casos, a soma de um par de elementos do conjunto e/ou a multiplicação de qualquer elemento por um escalar resulta em outro elemento do mesmo conjunto.

Para continuar, vejamos se conseguimos encontrar outras propriedades em comum que poderiam ser aplicadas a esses conjuntos.

Vetores no Plano - Lembrando do estudado na disciplina Geometria Analítica, não será difícil entender o significado das seguintes igualdades, (aqui $\alpha, \beta \in \mathbb{R}$ e $\mathbf{x}, \mathbf{y}, \mathbf{z}$ são vetores no plano).

1) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$;

2) $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$;

3) se $\mathbf{0}$ representa o vetor nulo, então $\mathbf{x}+\mathbf{0}=\mathbf{0}$;

4) o vetor $-\mathbf{x}$ (o oposto de $\mathbf{x}$) é tal que $\mathbf{x}+(-\mathbf{x})=\mathbf{0}$;

5) $\alpha(\beta\mathbf{x})=(\alpha\beta)\mathbf{x}$;

6) $(\alpha+\beta)\mathbf{x}=\alpha\mathbf{x}+\beta\mathbf{x}$;

7) $\alpha(\mathbf{x}+\mathbf{y})=\alpha\mathbf{x}+\alpha\mathbf{y}$;

8) $1\cdot\mathbf{x}=\mathbf{x}$.

Sugestão. Reproduza estes itens fazendo um esboço dos vetores resultantes em ambos os lados das igualdades e formule as propriedades algébricas correspondentes.

Funções reais - Verifica-se facilmente, a partir das propriedades dos números reais, que, com relação a quaisquer funções f, g e h em $F(\mathbb{R})$ e para todo $\alpha,\beta\in\mathbb{R}$, são válidos os seguintes resultados:

1) $f+g=g+f$;

2) $f+(g+h)=(f+g)+h$;

3) se 0 representa a função nula, isto é, $0(x)=0$ para todo x, então $0+f=0$;

4) a função $-f$ definida por $(-f(x))=-f(x)$ para todo $x\in\mathbb{R}$ é tal que $(-f)+f=0$;

5) $(\alpha f)=(\alpha)f$;

6) $(\alpha+\beta)f=\alpha f+\beta f$;

7) $\alpha(f+g)=\alpha f+\alpha g$;

8) $1\cdot f=f$.

Matrizes - Agora, com relação a quaisquer matrizes A, B e C em $\mathrm{M}_{n\times n}$ e para todo $\alpha,\beta\in\mathbb{R}$, também são válidos os seguintes resultados:

1) $A+B=B+A$;

2) $A+(B+C)=(A+B)+C$;

3) se O representa o função nula, isto é, $O=(0_{ij})_{n\times n}$, então $A+O=A$;

4) se $A=(a_{ij})_{n\times n}$, então a matriz $-A$ definida por $-A=(-a_{ij})_{n\times n}$ é tal que $A+(-A)=O$;

5) $\alpha(\beta A)=(\alpha\beta)A$;

6) $(\alpha+\beta)A=\alpha A+\beta A$;

7) $\alpha(A+B)=\alpha A+\alpha B$;

8) $1\cdot A=A$.

Observamos que os conjuntos considerados, quando munidos das operações de soma e multiplicação por escalares, apresentam propriedades algébricas comuns.

Na verdade, muitos outros conjuntos munidos das operações, de soma e multiplicação por escalar, apropriadas apresentam propriedades semelhantes às anteriores.

Em vez de investigar as propriedades de cada conjunto em particular, consideraremos um conjunto qualquer (desconsiderando a natureza de seus elementos) não vazio, V (cujos elementos chamaremos genericamente de **vetores**), sobre o qual supomos estarem definidas uma operação de adição, isto é, para cada $\mathbf{u},\mathbf{v}\in V$, existe um único elemento de V associado, chamado a soma entre $\mathbf{u}$ e $\mathbf{v}$ e denotado por $\mathbf{u}+\mathbf{v}$, e uma multiplicação por escalar, isto é, para cada $\mathbf{u}\in V$ e $\alpha\in\mathbb{R}$, existe um único vetor associado, chamado de o produto de $\mathbf{u}$ pelo escalar α e denotado por $\alpha\mathbf{u}$.

Estamos agora em condições de introduzir o conceito de espaço vetorial.

3.2 Espaços Vetoriais

3.2.1 Definição

Diremos que um conjunto V, munido de duas operações: uma de adição entre quaisquer elementos do comjunto e outra de multiplicação de quaisquer elemento do conjunto por escalar, é um *espaço vetorial*, se para quaisquer $\mathbf{u},\mathbf{v},\mathbf{w}\in V$, e para todo escalar $\alpha,\beta\in\mathbb{R}$, são válidos os seguintes axiomas:

O termo axioma é originado da palavra grega αξιωμα (axioma), que significa algo que é considerado, ajustado ou adequado, ou que tem um significado evidente. Entre os filósofos dos gregos antigos, um axioma era uma reivindicação que podia ser vista para ser verdade sem nenhuma necessidade de prova.

Existe a operação de adição: $\mathbf{u}+\mathbf{v}$ está em V (fechado sob adição). Isto significa que a soma de dois elementos quaisquer de V é um único elemento em V.

EV1) $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$ para quaisquer $\mathbf{u},\mathbf{v}\in V$. A soma é comutativa em V.

EV2) $\mathbf{u}+(\mathbf{v}+\mathbf{w})=(\mathbf{u}+\mathbf{v})+\mathbf{w}$ para quaisquer $\mathbf{u},\mathbf{v},\mathbf{w}\in V$. A soma é associativa em V.

EV3) Existe um elemento $\mathbf{0}\in V$, chamado vetor nulo tal que $0+\mathbf{u}=\mathbf{u}$ para todo $\mathbf{u}\in V$.

EV4) Para cada $\mathbf{u}\in V$ existe $\mathbf{-u}\in V$ tal que $\mathbf{u}+(-\mathbf{u})=0$; todo elemento de V admite o inverso aditivo ou oposto em V.

Existe a operação de multiplicação por escalar: $\alpha\mathbf{u}$ está em V (fechado sob multiplicação por escalar).

EV5) $\alpha(\beta\mathbf{u})=(\alpha\beta)\mathbf{u}$ para quaisquer $\mathbf{u}\in V$ e $\alpha,\beta\in\mathbb{R}$.

EV6) $(\alpha+\beta)\mathbf{u}=\alpha\mathbf{u}+\beta\mathbf{u}$ para quaisquer $\mathbf{u}\in V$ e $\alpha,\beta\in\mathbb{R}$.

EV7) $\alpha(\mathbf{u}+\mathbf{v})=\alpha\mathbf{u}+\alpha\mathbf{v}$ para quaisquer $\mathbf{u},\mathbf{v}\in V$ e $\alpha\in\mathbb{R}$; distributividade em relação à soma em V.

EV9) $1\mathbf{u}=\mathbf{u}$ para qualquer $\mathbf{u}\in V$.

Observações:

1) Os elementos de V são chamados de *vetores*.

2) Na maioria dos casos, os *escalares* serão os números reais. Dessa maneira, nos referiremos a V como um espaço vetorial real (ou espaço vetorial sobre os números reais). É possível também tomarmos os escalares como números complexos ou pertencentes a $\mathbb{Z}$. Os nossos exemplos serão de espaços vetoriais reais. Os escalares podem ser tomados de qualquer sistema numérico no qual podemos somar, subtrair, multiplicar e dividir de acordo com as leis habituais da aritmética.

3) A definição de espaço vetorial não especifica quais elementos formam o conjunto V. Também não especifica o que as operações chamadas "adição" e "multiplicação por escalar" devem ser.

4) O elemento $\mathbf{0}$, o vetor nulo , na propriedade EV3 é único, pois se existisse qualquer outro $0' \in V$ satisfazendo a EV3, teríamos:

$\mathbf{0}' \overset{\text{(por EV3)}}{=} \mathbf{0}' + \mathbf{0} \overset{\text{(por EV1)}}{=} \mathbf{0} + \mathbf{0}' = \mathbf{0}$, isto é, $\mathbf{0} = \mathbf{0}'$.

5) $\alpha + \beta$ e $\alpha\beta$ é soma e produto de escalares, respectivamente; $\mathbf{u} + \mathbf{v}$ é a soma de dois vetores em V; $\alpha\mathbf{u}$ é o produto de um escalar por um vetor.

Outro exemplo de espaço vetorial, além dos apresentados no início do texto, é o conjunto dos vetores de $\mathbb{R}^3$ (vetores do espaço tridimensional), como foram apresentados em Geometria Analítica, munidos da adição e da multiplicação por escalar.

O adjetivo vetorial utilizado na definição acima deve ser entendido de uma forma mais ampla, sendo uma referência aos elementos de V, independentemente de serem ou não vetores $\mathbb{R}^2$ ou $\mathbb{R}^3$.

3.2.2 Propriedades

Dos oito axiomas que definem um espaço vetorial, podemos deduzir várias outras propriedades. Alguma delas são relacionadas a seguir:

Seja $\mathbf{u} \in V$ um espaço vetorial. Temos que:

1) O produto de qualquer escalar pelo vetor nulo é o vetor nulo, isto é, para qualquer $\alpha \in \mathbb{R}$, $\alpha 0 = 0$.

2) O produto do escalar 0 por qualquer vetor é o vetor nulo, isto é, para qualquer $\mathbf{u} \in V$, $0\mathbf{u} = 0$.

3) Se o produto de um escalar por um vetor é o vetor nulo, então o escalar é 0 ou o vetor é o vetor nulo, isto é, se $\alpha\mathbf{u} = 0$, então $\alpha = 0$ ou $\mathbf{u} = 0$.

4) O produto do oposto de um escalar por um vetor é igual ao oposto de seu produto, isto é, para quaisquer $\alpha \in \mathbb{R}$ e $\mathbf{u} \in V$, $(-\alpha)\mathbf{u} = -(\alpha\mathbf{u})$.

5) Para quaisquer $\alpha, \beta \in \mathbb{R}$ e $\mathbf{u} \in V$, $(\alpha - \beta)\mathbf{u} = \alpha\mathbf{u} - \beta\mathbf{u}$.

6) Para quaisquer $\alpha \in \mathbb{R}$ e $\mathbf{u}, \mathbf{v} \in V$, $\alpha(\mathbf{u} - \mathbf{v}) = \alpha\mathbf{u} - \alpha\mathbf{v}$.

7) Para quaisquer $\alpha, \beta_1, \beta_2, \ldots, \beta_n \in \mathbb{R}$ e $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n \in V$

$$\alpha\left(\sum_{j=1}^{n} \beta_j \mathbf{u}_j\right) = \sum_{j=1}^{n} (\alpha\beta_j)\mathbf{u}_j .$$

8) Para qualquer $\mathbf{u} \in V$, $-(-\mathbf{u}) = \mathbf{u}$.

9) Se $\mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w}$, então $\mathbf{u} = \mathbf{v}$.

10) Se $\mathbf{u}, \mathbf{v} \in V$, então existe um único $\mathbf{w} \in V$ tal que $\mathbf{u} + \mathbf{w} = \mathbf{v}$.

A prova de algumas destas propriedades podem ser encontradas na bibliografia. Qualquer dúvida, procure o tutor de seu pólo!

3.2.3 Exemplos

Talvez o exemplo mais simples de espaço vetorial seja o conjunto dos números reais com a adição e multiplicação por escalar em forma usual.

Mas geralmente, para cada $n \in \mathbb{N}$, podemos transformar o conjunto das n-uplas ordenadas de números reais $\mathbb{R}^n$ (um vetor de $\mathbb{R}^n$ e uma matriz $n \times 1$ com componentes reais) em um espaço vetorial, definindo a adição de duas n-uplas ordenadas $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ e $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$ adicionando-se componente a componente, isto é, $\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)^T$ e o produto de uma n-upla $\mathbf{u} = (u_1, u_2, \ldots, u_n)^T$ por um escalar $\alpha \in \mathbb{R}$ por $\alpha\mathbf{u} = (\alpha u_1, \alpha u_2, \ldots \alpha u_n)^T$.

Exemplo 1. Consideremos $V = \mathbb{R}^n$ com a soma e a multiplicação por escalar, definidas como anteriormente (as usuais). Para verificarmos os oito axiomas de EV, consideremos $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$ e $\mathbf{z} = (z_1, z_2, \ldots, z_n)$. Tem-se que:

EV1) $$\begin{aligned}\mathbf{x} + \mathbf{y} &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)^T \\ &= (y_1 + x_1, y_2 + x_2, \ldots, y_n + x_n)^T \\ &= \mathbf{y} + \mathbf{x}\end{aligned}$$

EV2) $\mathbf{x}+(\mathbf{y}+\mathbf{z})=(x_1+(y_1+z_1),x_2+(y_2+z_2),\ldots,x_n+(y_n+z_n))^T$

$$=(x_1+y_1+z_1,x_2+y_2+z_2,\ldots,x_n+y_n+z_n)^T$$
$$=((x_1+y_1)+z_1,(x_2+y_2)+z_2,\ldots,(x_n+y_n)+z_n)^T$$
$$=(\mathbf{x}+\mathbf{y})+\mathbf{z}$$

EV3) $\mathbf{0}+\mathbf{x}=(0,0,\ldots,0)^T+(x_1,x_2,\ldots,x_n)^T$

$$=(0+x_1,0+x_2,\ldots,0+x_n)^T$$
$$=(x_1,x_2,\ldots,x_n)^T$$
$$=\mathbf{x}$$

EV4) $\mathbf{x}+(-\mathbf{x})=(x_1,x_2,\ldots,x_n)^T+(-(x_1,x_2,\ldots x_n)^T)$

$$=(x_1,x_2,\ldots,x_n)^T+(-x_1,-x_2,\ldots,-x_n)^T$$
$$=(x_1-x_1,x_2-x_2,\ldots,x_n-x_n)^T$$
$$=(0,0,\ldots0)^T$$
$$=\mathbf{0}$$

EV5) $\alpha(\beta\mathbf{x})=\alpha(\beta(x_1,x_2,\ldots,x_n)^T)$

$$=\alpha((\beta x_1,\beta x_2,\ldots,\beta x_n)^T)$$
$$=(\alpha\beta x_1,\alpha\beta x_2,\ldots,\alpha\beta x_n)^T$$
$$=\alpha\beta(x_1,x_2,\ldots,x_n)^T$$
$$=(\alpha\beta)\mathbf{x}$$

EV6) $(\alpha+\beta)\mathbf{x}=(\alpha+\beta)(x_1,x_2,\ldots,x_n)^T$

$$=((\alpha+\beta)x_1,(\alpha+\beta)x_2,\ldots,(\alpha+\beta)x_n)^T$$
$$=((\alpha x_1+\beta x_1),(\alpha x_2+\beta x_2),\ldots,(\alpha x_n+\beta x_n))^T$$
$$=(\alpha x_1,\alpha x_2,\ldots,\alpha x_n)^T+(\beta x_1,\beta x_2,\ldots,\beta x_n)^T$$
$$=\alpha\mathbf{x}+\beta\mathbf{x}$$

EV7) $\alpha(\mathbf{x}+\mathbf{y})=\alpha((x_1,x_2,\ldots,x_n)^T+(y_1,y_2,\ldots y_n)^T)$

$$=\alpha(x_1+y_1,x_2+y_2,\ldots,x_n+y_n)^T$$
$$=(\alpha(x_1+y_1),\alpha(x_2+y_2),\ldots,\alpha(x_n+y_n))^T$$
$$=\alpha\mathbf{x}+\alpha\mathbf{y}$$

EV8) $1\cdot\mathbf{x}=1\cdot(x_1,x_2,\ldots x_n)^T$

$$=(1\cdot x_1,1\cdot x_2,\ldots 1\cdot x_n)^T$$
$$=(x_1,x_2,\ldots,x_n)^T$$
$$=\mathbf{x}$$

Nos seguintes exemplos deve-se conferir que os conjuntos são espaços vetoriais.

Exemplo 2. Seja $V = P_2(\mathbb{R})$ o conjunto formado por todos os polinômios de grau menor ou igual a 2 com coeficientes reais. Definimos a adição e a multiplicação por escalar da seguinte maneira:

- Se $p(x) = a_0 + a_1x + a_2x^2$ e $q(x) = b_0 + b_1x + b_2x^2$ estão em P_2, então $p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2$ é outro polinômio que tem grau no máximo 2 e, portanto, está em P_2. Aqui dizemos que a soma entre vetores, ou a soma vetorial é uma operação fechada.
- Se c é um escalar, então $cp(x) = ca_0 + ca_1x + ca_2x^2$ também está em P_2. Também o produto por escalar é uma operação fechada.

Os dois primeiros axiomas são fáceis de verificar. Proceda como foi feito no exemplo anterior.

Você saberia dizer qual é o vetor nulo em P_2?

Existe um polinômio cujo valor seja sempre igual a 0 para todo x? O vetor nulo 0 é o polinômio zero, isto é, o polinômio cujos coeficientes são todos nulos: $a_0 = 0$, $a_1 = 0$ e $a_2 = 0$.

Confiramos EV4: o oposto de um polinômio $p(x) = a_0 + a_1x + a_2x^2$ é o polinômio $-p(x) = -(a_0 + a_1x + a_2x^2) = -a_0 - a_1x - a_2x^2$.

A satisfação dos axiomas EV5 ao EV8 se segue das propriedades dos números reais.

Observe que, se mudarmos o enunciado neste exemplo para: "Seja $V = P_2(\mathbb{R})$ o conjunto formado por todos os polinômios de grau igual a 2 (e não menor ou igual a 2) com coeficientes reais", não temos um espaço vetorial. Quais axiomas não são satisfeitos?

Confira se os conjuntos dos seguintes exemplos são espaços vetoriais.

Exemplo 3. O conjunto das funções contínuas definidas num intervalo $I \subset \mathbb{R}$ munido das operações de adição e multiplicação usuais.

Exemplo 4. O conjunto das matrizes $m \times n$ com coeficientes reais: $M_{m\times n}(\mathbb{R})$ munido de operações análogas àquelas definidas para matrizes quadradas.

Os espaços vetoriais acima citados envolvem operações com as quais você já deve estar familiarizado. O próximo exemplo é um pouco mais sofisticado do que os anteriores, no sentido de que definiremos as operações de adição e multiplicação por escalar em uma forma um pouco artificial (porém, que pode ser apropriada em situações particulares), e por isso mostraremos as oito propriedades.

Exemplo 5. Como conjunto objeto de estudo, escolheremos o intervalo semi-infinito $V = (0, \infty)$, (o semi-eixo positivo da reta real). Este conjunto, quando agregado às operações de soma e multiplicação usuais, não é um espaço vetorial, visto que não possui elemento neutro para a adição: $0 \notin V$ (0 não pertence ao conjunto), logo, EV3 não é satisfeita. No entanto, se para $\mathbf{x}, \mathbf{y} \in V$ e $\alpha \in \mathbb{R}$, definirmos a soma entre $\mathbf{x}$ e $\mathbf{y}$ por $\mathbf{x} \oplus \mathbf{y} = \mathbf{xy}$ (o produto usual entre $\mathbf{x}$ e $\mathbf{y}$) e o produto de $\mathbf{x}$ pelo escalar α como $\alpha \odot \mathbf{x} = \mathbf{x}$, então V se torna um espaço vetorial. De fato, verifiquemos uma a uma as oito propriedades:

EV1) temos $\mathbf{x} \oplus \mathbf{y} = \mathbf{xy} = \mathbf{yx} = \mathbf{y} \oplus \mathbf{x}$ para quaisquer $\mathbf{x}, \mathbf{y} \in V$;

EV2) $\mathbf{x} \oplus (\mathbf{y} \oplus \mathbf{z}) = \mathbf{x}(\mathbf{yz}) = (\mathbf{xy})\mathbf{z} = (\mathbf{x} \oplus \mathbf{y}) \oplus \mathbf{z}$ para quaisquer $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$;

EV3) se $\mathbf{x} \in V$, então como $1 \in V$, temos $1 \oplus \mathbf{x} = 1\mathbf{x} = \mathbf{x}$; observe que, neste caso, 1 é o elemento neutro da adição, o qual denotaremos por o;

EV4) se $\mathbf{x} \in V$, isto é, $\mathbf{x} > 0$, então $\mathbf{x}^{-1} \in V$ e $\mathbf{x} \oplus \mathbf{x}^{-1} = \mathbf{xx}^{-1} = 1 = o$;

EV5) $\alpha \odot (\beta \odot \mathbf{x}) = \alpha \odot \mathbf{x}^{\beta} = (\mathbf{x}^{\beta})^{\alpha} = \mathbf{x}^{\beta\alpha} = \mathbf{x}^{\alpha\beta} = (\alpha\beta) \odot \mathbf{x}$ para quaisquer $\mathrm{x} \in V$ e $\alpha, \beta \in \mathbb{R}$;

EV6) $(\alpha + \beta) \odot \mathbf{x} = \mathbf{x}^{(\alpha+\beta)} = \mathbf{x}^{\alpha} \oplus \mathbf{x}^{\beta} = (\alpha \odot \mathbf{x}) \oplus (\beta \odot \mathbf{x})$ para quaisquer $\mathbf{x} \in V$ e $\alpha + \beta \in \mathbb{R}$;

EV7) $\alpha \odot (\mathbf{x} \oplus \mathbf{y}) = \alpha \odot (\mathbf{xy}) = (\mathbf{xy})^{\alpha}$

$$= \mathbf{x}^{\alpha}\mathbf{y}^{\alpha} = \mathbf{x}^{\alpha} \oplus \mathbf{y}^{\alpha} = (\alpha \odot \mathbf{x}) \oplus (\alpha \odot \mathbf{y})$$

para quaisquer $\mathbf{x}, \mathbf{y} \in V$ e $\alpha \in \mathbb{R}$;

EV8) $1\mathbf{x} = \mathbf{x}^{-1} = \mathbf{x}$ para qualquer $\mathbf{x} \in V$.

3.2.4 Uma Definição Mais Formal

Você deve lembrar-se de que no início deste capítulo deixamos a definição de espaço vetorial, que chamamos de *"um tanto formal ou rigorosa"* em forma incompleta. Como prometido, essa definição é colocada em forma íntegra.

Um conjunto V é um espaço vetorial sobre um corpo K, que denotamos como $(V, \mathrm{K}, +, *)$, se dadas duas operações: a soma vetorial definida em V, que denotamos como $\mathbf{v} + \mathbf{w}$ para todo $\mathbf{v}, \mathbf{w} \in V$, e o produto escalar em V, que denotamos como $\alpha\mathbf{v}$ para todo $\mathbf{v} \in V$ e $\alpha \in \mathrm{K}$, se cumpre as seguintes 10 propriedades (5 propriedades para a soma e 5 para o produto escalar).

Para a Soma

1) $\mathbf{v} + \mathbf{w} \in V$. A soma vetorial é uma operação fechada em V.
2) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$. Associatividade da soma vetorial em V.
3) Existe um elemento $0 \in V$, tal que para todo $\mathbf{v} \in V$, $\mathbf{v} + 0 = \mathbf{v}$. Existência do elemento neutro da soma vetorial em V.
4) Para todo $\mathbf{v} \in V$, existe um elemento $-\mathbf{v} \in V$, tal que $\mathbf{v} + (-\mathbf{v}) = 0$. Existência do elemento oposto respeito à soma vetorial em V.
5) $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$. Comutatividade da soma vetorial em V.

Para o Produto por Escalar

1) $\alpha\mathbf{v} \in V$. O produto por escalar é uma operação fechada em V.
2) $\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}$. Associatividade do produto entre escalares em V.
3) Se 1 denota o elemento neutro da multiplicação do campo escalar K, então $1\mathbf{v} = \mathbf{v}$. Neutralidade do **um** do campo escalar.
4) $\alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$. Distributividade com respeito à soma vetorial.

5) $(\alpha+\beta)\mathbf{v}=\alpha\mathbf{v}+\beta\mathbf{v}$. Distributividade com respeito à soma escalar.

Exercícios Propostos

1) Considere os vetores $x_1=(8,6)^T$ e $x_2=(4,-1)^T$ em $\mathbb{R}^2$.

   a) Encontre o comprimento de cada vetor.

   b) Seja $x_3=x_1+x_2$. Determine o comprimento de x_3. Qual a relação entre seu comprimento e a soma dos comprimentos de x_1 e x_2?

   c) Desenhe um gráfico ilustrando como x_3 pode ser construído geometricamente usando x_1 e x_2. Use esse gráfico para dar uma interpretação geométrica da sua resposta em (b).

2) Repita o exercício 1 para os vetores $x_1=(2,1)^T$ e $x_2=(6,3)^T$.

3) Seja C o conjunto dos números complexos. Defina a soma em C por $(a+bi)+(c+di)=(a+c)+(b+d)i$ e defina a multiplicação por um escalar por $\alpha(a+bi)=\alpha a+\alpha bi$, para todos os números reais α. Mostre que C é um espaço vetorial em relação a essas operações.

4) Mostre que $\mathbb{R}^{m\times n}$, com as operações usuais de soma e multiplicação por um escalar, satisfaz os oito axiomas de espaços vetoriais.

5) Mostre que $C[a,b]$, com as operações usuais de soma e multiplicação por um escalar, satisfaz os oito axiomas de espaços vetoriais.

6) Seja P o conjunto de todos os polinômios. Mostre que P, com as operações usuais de soma e multiplicação por um escalar para funções, forma um espaço vetorial.

7) Sejam $\mathbf{x}, \mathbf{y}$ e $\mathbf{z}$ vetores de um espaço vetorial V. Mostre que, se $x+y=x+z$, então $y=x$.

8) Seja S o conjunto de todos os pares ordenados de números reais. Defina a multiplicação por um escalar e a soma em S por

$$\alpha(x_1,x_2)=(\alpha x_1,\alpha x_2)$$
$$(x_1,x_2)\oplus(y_1,y_2)=(x_1+y_1,0),$$

usando o símbolo $\oplus$ para denotar a soma nesse sistema, para evitar confusão com a soma usual de $\mathrm{x}+\mathrm{y}$ de vetores linhas. Mostre que S, junto com a multiplicação usual por um escalar e a operação $\oplus$, não é um espaço vetorial. Quais dos oito axiomas não são válidos?

9) Seja V o conjunto de todos os pares ordenados de números reais com a soma definida por $(x_1, x_2)+(y_1, y_2)=(x_1+y_1, x_2+y_2)$ e a multiplicação por um escalar definida por $\alpha \circ (x_1, x_2)=(\alpha x_1, x_2)$. Como a multiplicação por um escalar é definida de maneira diferente da usual, usamos um símbolo diferente para evitar confusão com a multiplicação usual de um vetor linha por um escalar. V é um espaço vetorial em relação a essas operações? Justifique sua resposta.

10) Denote por $\mathbb{R}^+$ o conjunto dos números reais positivos. Defina a operação de multiplicação por um escalar por $\alpha \circ x = x^\alpha$ para cada $x \in \mathbb{R}^+$ e para cada número real α. Defina a operação de soma por $x \oplus y = x \cdot y$ para todos $x, y \in \mathbb{R}^+$.

    Então, para esse sistema, o produto do escalar -3 por $\frac{1}{2}$ é dado por $-3 \circ \frac{1}{2} = \left(\frac{1}{2}\right)^{-3} = 8$ e a soma de 2 com 5 é dada por $2 \oplus 5 = 2 \cdot 5 = 10$. $\mathbb{R}^+$ é um espaço vetorial em relação a essas operações? Justifique sua resposta.

3.3 Subespaços Vetoriais

Atenção: aqui mudamos levemente a notação introduzida num exemplo no inicio deste capítulo, e, que um ponto de $\mathbb{R}^2$ foi denotado por (x_1, x_2).

Considere o espaço vetorial $\mathbb{R}^2$ formado por todos os pontos (x, y), com $x, y \in \mathbb{R}$.

À guisa de exemplo, consideremos dois subconjuntos de $\mathbb{R}^2$: uma reta que passa pela origem e o primeiro quadrante.

Reta que passa pela origem. Definimos L como sendo um subconjunto de $\mathbb{R}^2$ formado pelos pontos de uma reta que passa pela origem $(0,0)$, com equação $\alpha x + \beta y = 0$, ou seja, $L = \{(x, y) \in \mathbb{R}^2 \,/\, \alpha x + \beta y = 0\}$.

Observe que se $(x, y)(u, v) \in L$ (dois pontos quaisquer da reta), temos que $\alpha x + \beta y = 0$ e $\alpha u + \beta v = 0$, então $\alpha(x+u) + \beta(y+v) = 0$, desde que $(x, y) + (u, v) = ((x+u), (y+v)) \in L$. Portanto, a soma é fechada.

Vamos para os axiomas de espaço vetorial.

Note, primeiro de tudo, que $0 = (0,0) \in L$; assim, os axiomas do nulo e oposto são claramente satisfeitos em L. Também é fácil ver os de associatividade e comutatividade em relação à soma. Agora, se $(x, y) \in L$, então $\alpha x + \beta y = 0$ e $(cx, cy) \in L$ também, desde que $\alpha(cx) + \beta(cy) = 0$. Podemos concluir que L satisfaz todos os axiomas de espaço vetorial? Confira. De fato, qualquer reta que passe pela origem é um espaço vetorial sobre $\mathbb{R}^2$.

O que fizemos neste exemplo foi "extrair" de um espaço vetorial $\mathbb{R}^2$ outro espaço vetorial L "menor" e dizemos: L é um subespaço de $\mathbb{R}^2$. Logo, L é fechado em relação à soma e à multiplicação por escalar, isto é, a soma de dois elementos em L é um elemento de L e a multiplicação de um elemento de L por um escalar também pertence a L.

O primeiro quadrante. Quando analisamos o conjunto do segundo caso, o procedimento aparece equivalente ao anterior, no sentido de "extrair" um subconjunto de $\mathbb{R}^2$. Porém, agora nós afirmamos que este conjunto $Q = \{(x, y) \in \mathbb{R}^2, x, y \geq 0\}$ não é um espaço vetorial. Qual axioma não é satisfeito?

Na próxima seção introduziremos um outro conceito que nos permitirá entendermos quando um subconjunto é um espaço vetorial.

Esta seção introduz uma das idéias mais importantes desta disciplina. Começamos com um exemplo, vendo que há uma interação entre geometria e álgebra e que freqüentemente pode ser usada a intuição geométrica e lógica para introduzir propriedades e resolver problemas. Agora, é necessário tornar-nos mais formais, estendendo este conceito para *espaços vetoriais gerais.* A noção de subespaço é simplesmente uma generalização algébrica de retas e planos que passam pela origem.

Dado o espaço vetorial V e o conjunto não vazio $W \subset V$, se W é um espaço vetorial sobre o mesmo corpo K e com as mesmas leis de composição que em V, diremos que W é um subespaço de V.

3.3.1 Definição

Seja V um espaço vetorial. Dizemos que $W \subset V$, não vazio, é um *subespaço vetorial* de V se forem satisfeitas as seguintes condições:

SE1) $\mathbf{0} \in W$, o vetor nulo está em W.

SE2) Se $\mathbf{u}, \mathbf{v} \in W$, então $\mathbf{u} + \mathbf{v} \in W$.

SE3) Se $\mathbf{u} \in W$, então $\alpha \mathbf{u} \in W$ para todo $\alpha \in \mathbb{R}$.

Alternativamente, poderíamos dizer também que W é um subespaço de V se, e somente se, W é um espaço vetorial.

Observações:

1) Note que todo subespaço vetorial W de um espaço vetorial V é, ele próprio, um espaço vetorial.
2) O conjunto que consiste apenas no vetor nulo $\{0\}$ e o próprio V são chamados de subespaços vetoriais triviais.

Note que SE2 e SE3 podem ser reescritas equivalentemente como: se $\mathbf{u}, \mathbf{v} \in W$ e $\alpha \in \mathbb{R}$, então $\mathbf{u} + \alpha \mathbf{v} \in W$. Assim, as condições para ser subespaço podem ser resumidas em:

i) $\mathbf{0} \in W$, o vetor nulo está em W.

ii) Se $\mathbf{u}, \mathbf{v} \in W$ e $\alpha \in \mathbb{R}$ então $\mathbf{u} + \alpha \mathbf{v} \in W$.

Comentário. De fato, esta definição de subespaço vetorial é um resultado bastante útil para mostrar que um conjunto é um espaço vetorial.

3) Observe que SE3 implica diretamente que $W \neq \varnothing$ (W não é vazio).

Portanto, todo subespaço é um espaço vetorial e todo espaço vetorial é um subespaço (dele mesmo e, possivelmente, de outros espaços vetoriais maiores). O termo subespaço é usado quando pelo menos dois espaços vetoriais estão sendo considerados, um deles incluído no outro.

Desse modo, o "problema dos subespaços" consiste em determinar quando um subconjunto de um espaço vetorial continua satisfazendo todos os axiomas de espaço vetorial. O procedimento padrão para resolver este problema utiliza simplesmente a verificação das regras , dadas anteriormente.

Assim, na prática você deve responder as seguintes questões:

i) o vetor nulo de V está em W?

ii) W é fechado em relação a soma de vetores? Isto é, para cada $\mathbf{u}$ e $\mathbf{v} \in W$, a soma $\mathbf{u} + \mathbf{v}$ está em W?

iii) W é fechado em relação a multiplicação por escalar? Isto é, para cada $\mathbf{v} \in W$ e cada escalar α, o vetor $\alpha\mathbf{v}$ está em W?

3.3.2 Exemplos

Vejamos alguns exemplos:

Exemplo 6. Seja o espaço vetorial $V = \mathbb{R}^3$ e o conjunto W formado pelas triplas ordenadas de $V = \mathbb{R}^3$, tais que a terceira componente é igual à soma das duas primeiras, isto é, $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 / x_3 = x_2 + x_1\}$.

Afirmamos que W é um subespaço de V, desde que:

i) O vetor nulo $\mathbf{x} = (x_1, x_2, x_3) = (0,0,0)$ esteja em W:

$\mathbf{x} = (x_1, x_2, x_3) = (0,0,0) \in \mathbb{R}^3$ e a condição $x_3 = x_2 + x_1$ seja satisfeita para este vetor $(0 = 0 + 0)$.

ii) Sejam $\mathbf{x} = (x_1, x_2, x_3)$ e $\mathbf{y} = (y_1, y_2, y_3)$ vetores de W, tais que $x_3 = x_2 + x_1$ e $y_3 = y_2 + y_1$; ou seja, $\mathbf{x} = (x_1, x_2, x_2 + x_1)$ e $\mathbf{y} = (y_1, y_2, y_2 + y_1)$; com $\alpha \in \mathbb{R}$ calculamos $\mathbf{x} + \alpha\mathbf{y}$ e temos que

$$\begin{aligned}\mathbf{x}+\alpha\mathbf{y} &= (x_1, x_2, x_2 + x_1) + \alpha(y_1, y_2, y_2 + y_1) \\ &= (x_1 + \alpha y_1, x_2 + \alpha y_2, x_2 + \alpha y_2 + x_1 + \alpha y_1) \\ &= (z_1, z_2, z_2 + z_1) \in W.\end{aligned}$$

Então W é subespaço e a equação $x_1 + x_2 - x_3 = 0$ define um plano que passa pela origem.

Exemplo 7. Seja S o conjunto de todos os polinômios de grau menor ou igual a n, com a propriedade $p(0) = 0$. Verifiquemos se S é, de fato, um subespaço vetorial de P_n.

1) Se $a_j = 0, j = 1$ (todos os coeficientes nulos), então $p(x) = 0$ e $p(0) = 0$.
2) Se α é um escalar e $p(x) \in S$, então $\alpha p(0) = \alpha 0 = 0$ e $\alpha p \in S$.
3) Se $p(x) \in S$ e $q(x) \in S$, então $(p+q)(0) = p(0) + q(0) = 0 + 0 = 0$ e, portanto, $p + q \in S$.

Exemplo 8. Em qualquer espaço vetorial V, o vetor nulo forma o subespaço zero $\{0\}$.

Exemplo 9. $C(\mathbb{R})$ (o espaço das funções contínuas) é um subespaço de $F(\mathbb{R})$ porque a adição e multiplicação por escalar de funções contínuas são ainda contínuas. Em geral, $C^k(\mathbb{R})$ (o espaço das funções com k-ésima derivada contínua) é também um subespaço de $F(\mathbb{R})$.

Exemplo 10. O subconjunto $H = \{f \in F(\mathbb{R}) : f(2) = 0\}$ é também um subespaço de $F(\mathbb{R})$; se $f(2) = g(2) = 0$, então:

$$(f+g)(2) = f(2) + g(2) = 0 \text{ e } (cf)(2) = c(f(2)) = 0.$$

Exemplo 11. O subconjunto $K = \{f \in F(\mathbb{R}) : f(t) = f(-t)\}$ das funções pares é também um subespaço de $F(\mathbb{R})$; se $f(t) = f(-t)$ e $g(t) = g(-t)$, então:

$$(f+g)(t) = f(t) + g(t) = f(-t) + g(-t) = (f+g)(-t) \text{ e}$$
$$(cf)(t) = c(f(t)) = c(f(-t)) = cf(-t).$$

Por outro lado, o subconjunto $H = \{f \in F(\mathbb{R}) : f(2) = 1\}$ não é um subespaço de $F(\mathbb{R})$ pois não contém a função nula, que é o vetor zero do espaço vetorial $F(\mathbb{R})$.

Exemplo 12. P_3 é um subespaço de P_5. Por outro lado, H (conjunto de polinômios de grau maior ou igual a 3 e menor ou igual a 5) não é um subespaço de P_5. Um contra-exemplo servirá para mostrar que não é subespaço. Sejam $p(t) = 1 + t^4$ e $q(t) = -t^4$. Temos que $p, q \in H$, porém $(p+q)(t) = 1 \notin H$.

Exemplo 13. As soluções $T = \{(x_1, x_2, x_3) / x_1 + x_2 + x_3 = 0\}$ da equação homogênea são um subespaço de $\mathbb{R}^3$. Se (x_1, x_2, x_3) e $(y_1, y_2, y_3) \in T$, então $x_1 + x_2 + x_3 = 0$ e $y_1 + y_2 + y_3 = 0$; somando ambas igualdades, temos $(x_1 + y_1) + (x_2 + y_2) + (x_3 + y_3) = 0$, que satisfaz a equação homogênea e, então, pertencem a T. E também temos que

$$\alpha(x_1, x_2, x_3) = (\alpha x_1, \alpha x_2, \alpha x_3) \Rightarrow \alpha x_1 + \alpha x_2 + \alpha x_3 = (\alpha x_1) + (\alpha x_2) + (\alpha x_3) = 0$$

que satisfaz a equação homogênea e, portanto, pertence a T.

As soluções de $T = \{(x_1, x_2, x_3) / x_1 + x_2 + x_3 = 1\}$ (os pontos de um plano que não passam pela origem) da equação não-homogênea não é um subespaço de $\mathbb{R}^3$, já que não contém o zero $(0,0,0)$.

O primeiro quadrante $T = \{(x_1, x_2) / \ x_1 \geq 0, x_2 \geq 0\}$ de $\mathbb{R}^2$ não é um subespaço, desde que $(1,1) \in T$, porém $(-1)(1,1) = (-1,-1) \notin T$.

Os seguintes casos mostram outros conjuntos que não são subespaços de $\mathbb{R}^2$.

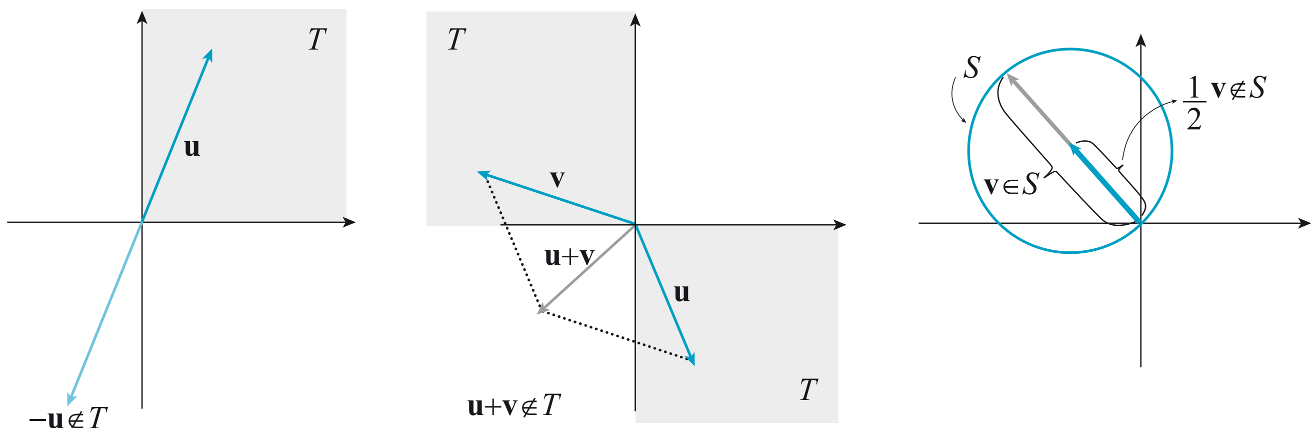


Figura 3.3 - Não são subespaços.

O gráfico seguinte mostra uma reta passando pela origem como sendo um subespaço; entretanto, uma reta que não passa pela origem não é um subespaço.

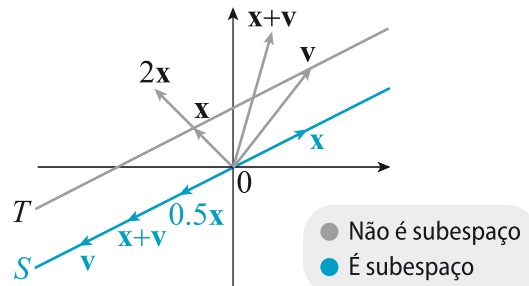


Figura 3.4 – Retas no plano.

Exemplo 14. $\mathbb{R}^2$ não é um subespaço de $\mathbb{R}^3$. Em verdade, $\mathbb{R}^2$ está definido por uma dupla de números reais, sendo que $\mathbb{R}^3$ são triplas de números, assim $\mathbb{R}^2$ não está contido em $\mathbb{R}^3$. Defina um plano coordenado do espaço tridimensional como sendo um subespaço de $\mathbb{R}^3$.

Exemplo 15. O conjunto $Sym(n)$ das matrizes simétricas é um subespaço do espaço vetorial $M_{2\times 2}$ das matrizes $n \times n$. A matriz que resulta da soma de matrizes simétricas é simétrica. Por outro lado, a matriz nula é simétrica e então pertence a $Sym(n)$.

Exemplo 16. Seja $V = M_{2\times 2}$ e $W = \{\text{matrizes inversíveis } 2\times 2\}$. Determine se W é um subespaço de $M_{2\times 2}$. Se W fosse um subespaço, a matriz nula deveria estar nele. Porém, $0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ não é inversível e, por isso, não pode estar em W. Logo, W não é um subespaço.

Em alguns dos nossos exemplos, vimos que subespaços aparecem como sendo exemplos de espaços vetoriais. **De fato, temos que um subespaço, com as operações herdadas da soma e multiplicação por escalar, é um espaço vetorial. A prova consiste em conferir as oito propriedades, a maior parte destas são herdadas de um espaço vetorial maior.**

3.3.3 Soma e Intersecção de Subespaços

Sejam H e K subespaços de um espaço vetorial V. Definimos a soma de dois subespaços como $H+K=\{\mathbf{u}+\mathbf{v} \,/\, \mathbf{u}\in H, \mathbf{v}\in K\}$ e a intersecção como $H\cap K=\{\mathbf{u} \,/\, \mathbf{u}\in H \wedge \mathbf{v}\in K\}$.

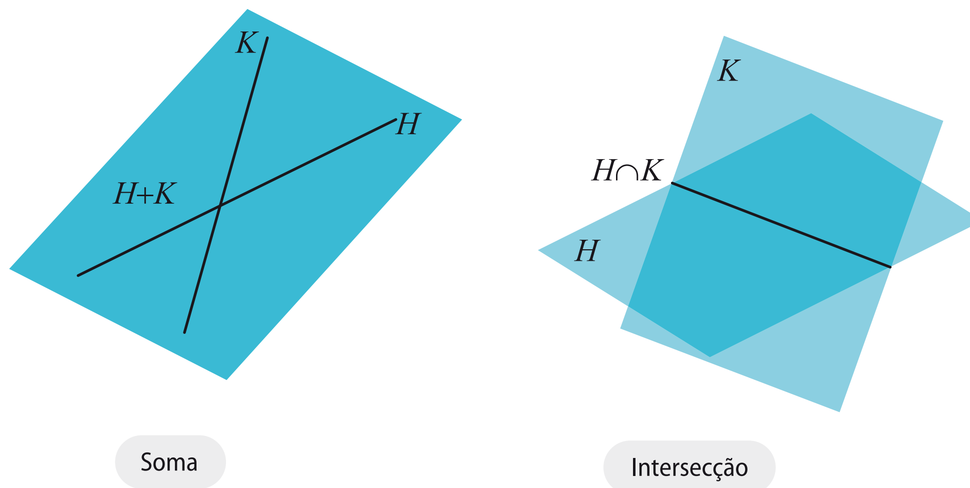


Figura 3.5 – Soma e intersecção de subespaços.

A soma e a intersecção de subespaços são subespaços vetoriais.

Observamos que a soma e a intersecção de espaços vetoriais são adaptações da união e intersecção de subconjuntos de espaços vetoriais. Enquanto a intersecção permanece sendo a mesma, a soma é diferente da união, já que, em geral, a união de dois subespaços não é um subespaço.

3.3.3.1 Intersecção de dois Subespaços Vetoriais

Sejam H e K dois subespaços vetoriais de V. A intersecção S de H e K, que se representa por $S=H\cap K$, é o conjunto de todos os vetores $\mathbf{v}\in V$ tais que $\mathbf{v}\in H$ e $\mathbf{v}\in K$.

Teorema

A intersecção S de dois subespaços vetoriais H e K de V é um subespaço vetorial de V. De fato:

1) Se $\mathbf{u}, \mathbf{v}\in H$, então $\mathbf{u}+\mathbf{v}\in H$
   Se $\mathbf{u}, \mathbf{v}\in K$, então $\mathbf{u}+\mathbf{v}\in K$
   Logo:
   $\mathbf{u}+\mathbf{v}\in S=H\cap K$

2) Para qualquer $\alpha \in \mathbb{R}$:
Se $\mathbf{v} \in H$, então $\alpha\ \mathbf{v} \in H$;
Se $\mathbf{v} \in K$, então $\alpha\ \mathbf{v} \in K$.
Logo:
$\alpha\ \mathbf{v} \in S = H \cap K$.

Exemplo 17. Seja V o espaço vetorial das matrizes quadradas de ordem 2:

$$V = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix}; a, b, c, d \in \mathbb{R} \right\}.$$

Considere:

$$H = \left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}; a, b \in \mathbb{R} \right\}$$

$$K = \left\{ \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix}; a, c \in \mathbb{R} \right\}.$$

Então, $\mathrm{S} = \mathrm{H} \cap \mathrm{K}$:

$$S = \left\{ \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}; a \in \mathbb{R} \right\}.$$

Exemplo 18. Seja o espaço vetorial $\mathbb{R}^3 = \{(\mathrm{a, b, c});\ \mathrm{a, b, c} \in \mathbb{R}\}$ e os subespaço $H = \{(\mathrm{a, b, c});\ \mathrm{a, b, c} \in \mathbb{R}\}$ e $K = \{(0, 0, \mathrm{c});\ \mathrm{c} \in \mathbb{R}\}$. A intersecção $H \cap K$ é o subespaço vetorial $S = \{(0, 0, 0)\} = \{0\}$.

3.3.3.2 Soma de dois subespaços vetoriais

Sejam H e K dois subespaços vetoriais de V. A soma S de H e K, que se representa por $S = H + K$, é o conjunto de todos os vetores $\mathbf{u} + \mathbf{v}$ de V tais que $\mathbf{u} \in H$ e $\mathbf{v} \in K$.

Teorema

A soma S de dois subespaços vetoriais H e K de V é um subespaço vetorial de V. De fato:

1) Se $\mathbf{u_1}, \mathbf{u_2} \in H$, então $\mathbf{u_1} + \mathbf{u_2} \in H$
Se $\mathbf{v_1}, \mathbf{v_2} \in K$, então $\mathbf{v_1} + \mathbf{v_2} \in K$

Por outro lado:

$$\mathbf{u}_1 + \mathbf{v}_1 \in S$$
$$\mathbf{u}_2 + \mathbf{v}_2 \in S.$$

Logo:

$$(\mathbf{u}_1 + \mathbf{v}_1) + (\mathbf{u}_2 + \mathbf{v}_2) = (\mathbf{u}_1 + \mathbf{u}_2) + (\mathbf{v}_1 + \mathbf{v}_2) \in H + K = S.$$

2) Para qualquer $\alpha \in \mathbb{R}$;

Se $\mathbf{u}_1 \in H$, então $\alpha\,\mathbf{u}_1 \in H$

Se $\mathbf{v}_1 \in K$, então $\alpha\,\mathbf{v}_1 \in K$.

Por outro lado:

$$\mathbf{u}_1 + \mathbf{v}_1 \in S.$$

Logo:

$$\alpha(\mathbf{u}_1 + \mathbf{v}_1) = \alpha\mathbf{u}_1 + \alpha\mathbf{v}_1 \in H + K = S.$$

Exemplo 19. Sejam os subespaços vetoriais $H = \{(a, b, 0); a, b \in \mathbb{R}\}$ e $K = \{(0, 0, c); c \in \mathbb{R}\}$ do espaço vetorial $\mathbb{R}^3 = \{(a, b, c); a, b, c \in \mathbb{R}\}$.

A soma $H + K$ é o subespaço vetorial $S = \{(a, b, c); a, b, c \in \mathbb{R}\}$, que no caso, é o próprio $\mathbb{R}^3$.

3.3.3.3 Soma direta de dois subespaços vetoriais

Sejam H e K dois subespaços vetoriais de V.

Diz-se que V é a soma direta de H e K e representa-se por $V = H \oplus K$. Se $V = H + K$ e $H \cap K = \{0\}$.

Isto é, se o único vetor comum a ambos os subespaços H e K for o vetor nulo.

Os símbolos $\odot$ e $\oplus$ são utilizados para indicar que a adição e a multiplicação por escalar não são as usuais.

Exemplo:

$$H = \{(a, 0, c, 0); a, c \in \mathbb{R}\}$$
$$K = \{(0, b, 0, d); b, d \in \mathbb{R}\}$$

Então:

$$\begin{cases} H + K = \{(a, b, c, d);\ a, b, c, d \in \mathbb{R}\} = \mathbb{R}^4 \\ H \cap K = \{(0, 0, 0, 0) = 0\} \end{cases}$$

Logo $H \oplus K$.

Exercícios Propostos

1) Considere os subconjuntos de $\mathbb{R}^3$:

$F = \{(x_1,\ x_2,\ x_3); x_1 + x_2 = 0 = x_3\ \}$ e $G = \{(x_1, x_2, x_3); x_1 + x_2 = 0\}$.

a) Prove que os conjuntos são subespaços vetoriais.

b) Diga, justificando, se a união dos conjuntos é um espaço vetorial.

c) Determine $F + G$ e $F \cap G$. Diga se a soma é direta.

2) Se um sistema linear não for homogêneo, o que acontece com seu conjunto solução? Considere o exemplo:

$$A = \begin{cases} 2x + 4y + z = 1 \\ x + y + 2z = 1 \\ x + 3y - z = 0 \end{cases}.$$

Provar que a soma de dois vetores solução nem sempre é um vetor solução, e assim o conjunto solução não é um subespaço vetorial.

A soma direta de subespaços tem conseqüências importantes, uma vez que nos permite “decompor” os espaços vetoriais em termos de subespaços, dos quais se espera que sejam de natureza mais simples. De fato, a maioria dos teoremas de decomposição tem suas raízes conceituais nesta idéia.

3.3.4 O espaço Nulo de A

Introduziremos agora um dos mais importantes subespaços de $\mathbb{R}^n$ associado a uma matriz.

Outros subespaços associados a uma matriz serão estudados em seções posteriores.

Suponhamos que A seja uma matriz $m \times n$. As soluções do sistema $A\mathbf{x} = 0$ determinam um subespaço de $\mathbb{R}^n$ chamado de *espaço nulo de* A, que denotaremos por $espnul(A) = \{x \in \mathbb{R}^n \ / \ A\mathbf{x} = 0\}$.

Consideremos agora alguns exemplos que são de fácil resolução.

Exemplo 20. Determine $espnul(A)$ para cada uma das seguintes matrizes.

Veremos outros exemplos de espaços nulos na seção de Bases de um espaço vetorial.

a) $A = \begin{pmatrix} 2 & 0 \\ -4 & 10 \end{pmatrix}$

b) $B = \begin{pmatrix} 1 & -7 \\ -3 & 21 \end{pmatrix}$

Solução. a) Para achar o espaço nulo de A, podemos escalonar a matriz ou resolver um sistema de duas equações com duas incógnitas. Neste caso, a última opção é a mais direta e temos que $x_1 = x_2 = 0$ é a solução. Em termos de vetores, a nossa solução consiste em um único vetor $\{0\}$, e então $espnul(A) = \{0\}$.

b) Aqui temos infinitas soluções da forma $x_1 = 7\alpha, x_2 = \alpha$, onde α é um escalar. Assim, já que o espaço nulo de B está formado por todas as soluções de $B\mathbf{x} = 0$ da forma $\mathbf{x} = (7\alpha, \alpha)^T = \alpha(7,1)^T$, sendo α um número real, então $espnul(B) = \left\{ \alpha \begin{pmatrix} 7 \\ 1 \end{pmatrix} \right\}$.

Veremos na próxima seção uma forma mais adequada de escrever este último resultado.

Podemos observar que o espaço nulo está descrito por todos os vetores associados aos pontos localizados na reta que passa pela origem com equação $x_1 - 7x_2 = 0$.

Voltemos agora aos nossos exemplos anteriores. O espaço nulo para a primeira matriz foi $\{0\}$. Para a segunda matriz, o espaço nulo foi uma reta passando pela origem. Lembrando os conceitos mencionados previamente, podemos ver que em ambos os casos, obtiveram-se subespaços de $\mathbb{R}^2$.

De fato, isso acontecerá sempre, como estabelece o teorema a seguir.

Teorema. *Seja* A *uma matriz* $m \times n$*, então o espaço nulo de* A *será um subespaço de* $\mathbb{R}^n$*.*

Prova. Sabemos que este subespaço está formado por todas as soluções do sistema homogêneo $A\mathbf{x} = 0$. Primeiro, deveríamos conferir que o vetor zero (0), que está em $\mathbb{R}^n$ é solução do sistema, e então deduziremos que o espaço nulo não é vazio, o que, neste caso, é trivial (qual é o produto que deve ser feito para mostrar isto?). Logo, ao saber que o espaço nulo tem, no mínimo, um vetor, escolheremos dois vetores $\mathbf{x}$ e $\mathbf{y}$ do espaço nulo e um escalar α, e conferiremos as leis de composição interna.

Começamos pela soma.

$$A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = 0 + 0 = 0$$

A soma $\mathbf{z} = \mathbf{x} + \mathbf{y}$ é a solução para $A\mathbf{z} = 0$, então a soma está no espaço nulo. Portanto, o espaço nulo é fechado sob a soma.

Logo, testamos a multiplicação por escalar.

$$A(\alpha\mathbf{x}) = \alpha A\mathbf{x} = \alpha 0 = 0$$

Se $\mathbf{x}$ está no espaço nulo, todo múltiplo de $\mathbf{x}$ também está e o espaço nulo é fechado sob multiplicação por escalar.

Então o espaço nulo é um subespaço de $\mathbb{R}^n$. Para verificar se um vetor de V está no espaço nulo de A, basta calcular $A\mathbf{x}$ e ver se $A\mathbf{x} = 0$.

Exercícios Propostos

1) Determine se cada conjunto a seguir é ou não um subespaço de $\mathbb{R}^2$.

a) $\{(x_1, x_2)^T / x_1 + x_2 = 0\}$

b) $\{(x_1, x_2)^T / x_1 \cdot x_2 = 0\}$

c) $\{(x_1, x_2)^T / x_1 = 3x_2\}$

d) $\{(x_1, x_2)^T / x_1 = 3x_2 + 1\}$

2) Determine se cada conjunto a seguir é ou não um subespaço de $\mathbb{R}^3$.

a) $\{(x_1, x_2, x_3)^T \,/\, x_1 + x_3 = 1\}$

b) $\{(x_1, x_2, x_3)^T \,/\, x_1 = x_2 = x_3\}$

c) $\{(x_1, x_2, x_3)^T \,/\, x_3 = x_1 + x_2\}$

d) $\{(x_1, x_2, x_3)^T \,/\, x_3 = x_1^{\,2} + x_2^{\,2}\}$

3) Determine se cada conjunto a seguir é ou não um subespaço de $\mathbb{R}^{2\times 2}$.

a) O conjunto de todas as matrizes diagonais 2×2.

b) O conjunto de todas as matrizes triangulares inferiores a 2×2.

c) O conjunto de todas as matrizes 2×2, tais que $a_{12} = 1$.

d) O conjunto de todas as matrizes 2×2, tais que $b_{11} = 0$.

e) O conjunto de todas as matrizes simétricas 2×2.

f) O conjunto de todas as matrizes singulares 2×2.

4) Determine o núcleo de cada uma das matrizes a seguir.

a) $\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$

b) $\begin{pmatrix} 1 & 2 & -3 & -1 \\ -2 & -4 & 6 & 3 \end{pmatrix}$

c) $\begin{pmatrix} 1 & 3 & -4 \\ 2 & -1 & -1 \\ -1 & -3 & 4 \end{pmatrix}$

d) $\begin{pmatrix} 1 & 1 & -1 & 2 \\ 2 & 2 & -3 & 1 \\ -1 & -1 & 0 & -5 \end{pmatrix}$

5) Determine se cada conjunto a seguir é ou não um subespaço de P4 . (Cuidado!)

a) O conjunto dos polinômios em P_4 de grau par.

b) O conjunto dos polinômios de grau 3.

c) O conjunto dos polinômios $p(x)$ em P_4, tais que $p(0) = 0$.

d) O conjunto dos polinômios em P_4 que tem pelo menos uma raiz real.

6) Determine se cada conjunto a seguir é ou não um subespaço de $C[-1,1]$.

a) O conjunto das funções f em $C[-1,1]$, tais $f(-1) = f(1)$.

b) O conjunto das funções ímpares em $C[-1,1]$.

c) O conjunto das funções não decrescentes em $[-1,1]$.

d) O conjunto das funções em f em $C[-1,1]$, tais $f(-1) = 0$ e $f(1) = 0$.

e) O conjunto das funções f em $C[-1,1]$, tais $f(-1) = 0$ ou $f(1) = 0$.

7) Determine se cada conjunto a seguir é ou não um conjunto gerador para $\mathbb{R}^2$.

a) $\left\{\begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}3\\2\end{pmatrix}\right\}$

b) $\left\{\begin{pmatrix}2\\3\end{pmatrix}, \begin{pmatrix}4\\6\end{pmatrix}\right\}$

c) $\left\{\begin{pmatrix}-2\\1\end{pmatrix}, \begin{pmatrix}1\\3\end{pmatrix}, \begin{pmatrix}2\\4\end{pmatrix}\right\}$

d) $\left\{\begin{pmatrix}-1\\2\end{pmatrix}, \begin{pmatrix}1\\-2\end{pmatrix}, \begin{pmatrix}2\\-4\end{pmatrix}\right\}$

e) $\left\{\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}\right\}$

8) Quais dos conjuntos a seguir são ou não conjuntos geradores para $\mathbb{R}^3$? Justifique suas respostas.

a) $\{(1,0,0)^T, (0,1,1)^T, (1,0,1)^T\}$

b) $\{(1,0,0)^T, (0,1,1)^T, (1,0,1)^T, (1,2,3)^T\}$

c) $\{(2,1,-2)^T, (3,2,-2)^T, (2,2,0)^T\}$

d) {(2 , 1 , -2)T, (-2 , -1 , 2)T, (4 , 2 , -4)T}

e) $\{(1,1,3)^T, (0,2,1)^T\}$

9) Sejam $x_1 = \begin{pmatrix}-1\\2\\3\end{pmatrix}$, $x_2 = \begin{pmatrix}3\\4\\2\end{pmatrix}$, $x = \begin{pmatrix}2\\6\\6\end{pmatrix}$, $y = \begin{pmatrix}-9\\-2\\5\end{pmatrix}$.

a) $x \in [\{x_1, x_2\}]$?

b) $y \in [\{x_1, x_2\}]$?

10) Quais dos conjuntos a seguir são conjuntos geradores para P_3? Justifique suas respostas.

a) $\{1, x^2, x^2 - 2\}$

b) $\{2, x^2, 2x + 3\}$

c) $\{x + 2, x + 1, x^2 - 1\}$

d) $\{x + 2x, x^2 - 1\}$

11) Em $\mathbb{R}^{2\times 2}$, sejam

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Mostre que $E_{11}, E_{12}, E_{21}, E_{22}$ geram $\mathbb{R}^{2\times 2}$.

3.4 Espaços Gerados

O tópico da seção anterior nos permite introduzir um conceito chave desta disciplina. Voltemos para o item (b) do exemplo 20. Neste exemplo, vimos que o espaço nulo está formado por todos os vetores da forma $\mathbf{x} = (7\alpha, \alpha)^T = \alpha(7,1)^T$ (onde α é um número real).

Para expressar este resultado em uma forma mais compacta e adequada, revisaremos um conceito que foi introduzido em um capítulo anterior. Quando foram estudadas as operações com matrizes, calculamos combinações lineares de linhas das matrizes e combinações lineares das colunas de uma matriz.

Podemos também considerar uma combinação linear de vetores.

Definição. Dizemos que o vetor $\mathbf{w}$ de um espaço vetorial V é uma combinação linear dos vetores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in V$ se existirem escalares $c_1, c_2, \ldots, c_n$, tal que $\mathbf{w}$ possa ser escrito como $\mathbf{w} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n}$.

Podemos observar que o espaço nulo mencionado acima pode ser visto, de fato, como todas as combinações lineares do vetor $(7,1)^T$. Até pode parecer estranho falar de uma combinação linear de um

único vetor, porém, de acordo com a definição, não há impedimento para considerá-lo desta forma.

Provavelmente, não é a primeira vez que tenha encontrado uma combinação linear de vetores. Na disciplina Geometria Analítica, foi introduzido o espaço euclidiano junto com o que chamamos de vetores canônicos. O conjunto de vetores canônicos para $\mathbb{R}^n$ foi definido como $\mathbf{e_1} = (1,0,\ldots,0)^T$, $\mathbf{e_2} = (0,1,\ldots,0)^T$, …, $\mathbf{e_n} = (0,0,\ldots,1)^T$.

Vimos que qualquer vetor $\mathbf{u} = (u_1, u_2, \ldots, u_n)^T \in \mathbb{R}^n$ pode ser escrito como $\mathbf{u} = u_1\mathbf{e_1} + u_2\mathbf{e_2} + \ldots + u_n\mathbf{e_n}$.

Em outras palavras, poderíamos ver o vetor $\mathbf{u}$ como uma combinação linear dos vetores da base canônica $\mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_n}$.

Exemplo 21. Determine, em cada caso, se o vetor $\mathbf{w}$ é uma combinação linear dos vetores $\mathbf{v_1}$ e $\mathbf{v_2}$.

a) $\mathbf{w} = (-12, 20)^T$, $\mathbf{v_1} = (-1, 2)^T$, $\mathbf{v_2} = (4, -6)^T$

b) $\mathbf{w} = (4, 20)^T$, $\mathbf{v_1} = (2, 10)^T$, $\mathbf{v_2} = (-3, -15)^T$

c) $\mathbf{w} = (1, -4)^T$, $\mathbf{v_1} = (2, 10)^T$, $\mathbf{v_2} = (-3, -15)^T$

Solução. Em cada um destes casos, precisamos resolver $\mathbf{w} = c_1\mathbf{v_1} + c_2\mathbf{v_2}$, isto é, achar c_1 e c_2.

a) $\mathbf{w} = (-12, 20)^T = c_1(-1, 2)^T + c_2(4, -6)^T \rightarrow \begin{cases} -c_1 + 4c_2 = -12 \\ 2c_1 - 6c_2 = 20 \end{cases}$.

Se o sistema for consistente (isto é, possuir no mínimo uma solução), então $\mathbf{w}$ é uma combinação linear de dois vetores; se não existir solução, então $\mathbf{w}$ não é uma combinação linear.

Deixamos para você verificar que a solução do sistema é $c_1 = 4$ e $c_2 = -2$.

Logo, $\mathbf{w}$ é uma combinação linear de $\mathbf{v_1}$ e $\mathbf{v}_2$ e pode ser escrito como $\mathbf{w} = 4\mathbf{v_1} - 2\mathbf{v_2}$.

b) Repetindo o procedimento que fizemos no item anterior e resolvendo o sistema, obtemos como solução $c_1 = 2 + \frac{3}{2}t$ e $c_2 = t$, onde t é um escalar. Isto significa que $\mathbf{w}$ é uma combinação linear de $\mathbf{v_1}$ e

$\mathbf{v}_2$ e pode ser escrita em um número infinito de diferentes combinações lineares. Algumas delas são, por exemplo:

$$\mathbf{w} = 2v_1 + (0)v_2, \mathbf{w} = 8v_1 + 4v_2, \mathbf{w} = (0)v_1 - \frac{4}{3}v_2, \mathbf{w} = v_1 - 2v_2.$$

c) Neste último caso, não temos solução, e então $\mathbf{w}$ não pode ser escrito como uma combinação linear de $\mathbf{v_1}$ e $\mathbf{v}_2$.

Embora estes exemplos sejam muito simples, o procedimento para um número maior de vetores continua sendo o mesmo.

Neste ponto, podemos dizer que sabemos como funciona uma combinação linear e responder quando um vetor é combinação linear de outro conjunto de vetores, e dessa forma, estamos em condições de avançar para o principal conceito desta seção.

Definição. Seja $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ um conjunto de vetores em um espaço vetorial V e seja W o conjunto de todas as combinações lineares dos vetores $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$. Dizemos que W é o espaço gerado pelos vetores $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ e escrevemos $W = \text{span}(S)$ ou $W = \text{span}\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ ou $W = [\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}]$. Também dizemos que $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ *gera* W.

Com esta notação, voltando ao exemplo do espaço nulo, temos simplesmente que $espnul(A) = \left[\begin{pmatrix} 7 \\ 1 \end{pmatrix}\right]$.

Temos o seguinte teorema:

Teorema. *Sejam* $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ *vetores de um espaço vetorial* V *e* $W = [\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}]$ *o espaço gerado, então temos que:*

a) W é um subespaço de V.

b) W é o menor subespaço de V que contém todos os vetores $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$.

Exemplo 22. Descreva o espaço gerado pelos seguintes conjuntos de vetores.

a) $\mathbf{v_1} = (1,0,0)^T, \mathbf{v_2} = (0,1,0)^T$

b) $\mathbf{v_1} = (1,0,1,0)^T, \mathbf{v_2} = (0,1,0,-1)^T$

c) $\mathbf{v_1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{v_2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

d) $\mathbf{v_1} = 1, \mathbf{v_2} = x, \mathbf{v_3} = x^3$

Solução. a) O espaço gerado, $[\mathrm{v}_1, \mathrm{v}_2]$ é o conjunto de todas as combinações lineares de $\mathbf{v_1}, \mathbf{v_2}$ que podem ser escritas como: $a\mathbf{v_1} + b\mathbf{v_2} = (a,0,0)^T + (0,b,0)^T$, sendo a e b escalares.

Logo, o espaço gerado são todos os vetores da forma $(a,b,0)^T \in \mathbb{R}^3$, sendo a e b escalares arbitrários.

b) Este caso é muito similar ao anterior. A combinação linear geral é: $a\mathrm{v}_1 + b\mathrm{v}_2 = (a,0,a,0)^T + (0,b,0,-b)^T = (a,b,a,-b)^T$.

Assim, $[\mathbf{v_1}, \mathbf{v_2}]$ são todos os vetores de $\mathbb{R}^4$ da forma $(a,b,a,-b)^T$ para quaisquer escalares a e b.

Atenção! Neste item, o procedimento é análogo aos itens a) e b), apenas com a diferença que não estamos trabalhando em $\mathbb{R}^n$.

c) Escrevemos a combinação linear destes vetores

$$a\mathbf{v_1} + b\mathbf{v_2} = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

O espaço gerado, $W = \mathrm{span}\{\mathbf{v_1}, \mathbf{v_2}\}$, é o subespaço das matrizes diagonais das matrizes $M_{2\times 2}$.

d) A combinação linear geral, neste caso, é $a\mathrm{v}_1 + b\mathrm{v}_2 + c\mathrm{v}_3 = a + bx + cx^3$. Aqui, $[\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}]$ é o subespaço dos polinômios de P_3 que não têm o termo quadrático.

Vejamos agora se conseguimos determinar o conjunto de vetores que geram os espaços dos exemplos anteriores, de tal forma que qualquer vetor possa ser escrito como uma combinação linear dos vetores do conjunto.

Exemplo 23. Determine um conjunto de vetores que gerem exatamente cada um dos seguintes espaços vetoriais.

a) $V = \mathbb{R}^n$

b) $V = M_{2\times 2}$

c) $V = P_n$

Pense no que realmente precisa ser verificado!

Solução. Necessitamos achar um conjunto de vetores, tal que o espaço gerado por este conjunto seja exatamente o espaço dado. Em outras palavras, devemos mostrar que o espaço gerado do nosso conjunto seja, de fato, o mesmo que o espaço vetorial dado.

O que deve ser feito para mostrar isso? Suponhamos que queremos mostrar que A e B são dois conjuntos iguais. Para isto, primeiro devemos mostrar que cada a em A também está em B, e então mostraremos que B, no mínimo, conterá totalmente o A. Analogamente, se cada b em B está em A, B está contido em A. Portanto, se temos que A contém B e B contém A, podemos concluir que A e B são o mesmo conjunto.

Assim, para os nossos exemplos, em cada caso, necessitamos exibir um conjunto de vetores geradores, tal que cada vetor do espaço vetorial dado V esteja nesse conjunto, para depois mostrar que cada vetor do conjunto gerador deve estar no espaço vetorial dado.

a) Já mostramos que cada vetor de $\mathbb{R}^n$ pode ser escrito como uma combinação linear dos vetores canônicos, $\mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_n}$, e assim os vetores canônicos geram um espaço que, no mínimo, contém todo $\mathbb{R}^n$. Por outro lado, já que qualquer combinação linear dos vetores canônicos deve ser um vetor de $\mathbb{R}^n$, podemos concluir que $\mathbb{R}^n$ deve conter o espaço gerado pelos vetores canônicos.

Logo, o espaço gerado pelo conjunto de vetores canônicos deve ser $\mathbb{R}^n$.

b) Em um exemplo anterior, vimos que um conjunto de matrizes gera o espaço das matrizes diagonais de $M_{2\times 2}$. Há uma extensão natural para que este conjunto gere o espaço de todas as matrizes 2×2, acrescentando mais dois vetores ao conjunto dado. De fato, o seguinte conjunto faz isso.

$$\mathbf{v_1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{v_2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \mathbf{v_3} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \mathbf{v_4} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Claramente, qualquer combinação linear destas quatro matrizes será uma matriz 2×2, e então o espaço gerado por estas matrizes deve estar contido em $M_{2\times 2}$.

Por outro lado, dada qualquer matriz de $M_{2\times 2}$, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2\times 2}$ pode ser escrita como uma combinação linear destes vetores. $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\mathbf{v_1} + b\mathbf{v_2} + c\mathbf{v_3} + d\mathbf{v_4}$, e assim, todo $M_{2\times 2}$ deve estar contido no espaço gerado por estes vetores, de tal forma que estes vetores geram $M_{2\times 2}$.

c) Lembremos que P_n é o conjunto de todos os polinômios de grau menor ou igual a n. Usando um dos exemplos anteriores como guia, estendemos o conjunto em forma natural para $\mathbf{v_1} = 1, \mathbf{v_2} = x, \mathbf{v_3} = x^2, \ldots, \mathbf{v_{n+1}} = x^n$. Uma combinação linear desses vetores é um polinômio de grau: n ou menor, e então estará em P_n. Logo, o espaço gerado estará em P_n. Agora, podemos escrever qualquer polinômio de grau n ou menor,

$$\text{p} = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

como a seguinte combinação linear:

$$\text{p} = a_0 \mathbf{v_1} + a_1 \mathbf{v_2} + a_2 \mathbf{v_3} + \ldots + a_n \mathbf{v_{n+1}}.$$

Então, todo P_n está contido no espaço gerado por estes vetores, o que significa que $[\{1, x, x^2, \ldots, x^n\}]$ é exatamente P_n.

Ainda é necessário discutir um último assunto sobre espaços gerados, e uma boa forma de fazer isso é através de um exemplo.

Exemplo 24. Determine se os seguintes conjuntos geram $\mathbb{R}^3$.

a) $\mathbf{v_1} = (2,0,1)^T, \mathbf{v_2} = (-1,3,4)^T, \mathbf{v_3}(1,1,-2)^T$

b) $\mathbf{v_1} = (1,2,-1)^T, \mathbf{v_2} = (3,-1,1)^T, \mathbf{v_3}(-3,8,-5)^T$

Solução. a) Claramente, o espaço gerado por esses vetores estará em $\mathbb{R}^3$. O problema agora é determinar se $\mathbb{R}^3$ estará contido no espaço gerado por esses vetores.

No exemplo anterior, conseguimos responder com facilidade a esta questão. Entretanto, neste caso, não parece ser tão evidente.

Para resolver este problema, consideremos o seguinte procedimento:

Escolhamos arbitrariamente um vetor de $\mathbb{R}^3$, $\mathbf{u} = (u_1, u_2, u_3)^T$, e tentemos encontrar escalares c_1, c_2, c_3, de tal forma que seja possível escrever u como uma combinação linear dos vetores dados $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$. Ou seja,

$$\mathbf{u} = (u_1, u_2, u_3) = c_1\mathbf{v_1} + c_2\mathbf{v_2} + c_3\mathbf{v_3} = c_1(2,0,1)^T + c_2(-1,3,4)^T + c_3(1,1,-2)^T.$$

Igualando os respectivos componentes, obtemos o seguinte sistema de equações:

$$\begin{aligned} 2c_1 - c_2 + c_3 &= u_1 \\ 3c_2 + c_3 &= u_2 \\ 3c_1 + 4c_2 - 2c_3 &= u_3 \end{aligned}$$

Que podemos escrever em forma matricial como:

$$\begin{pmatrix} 2 & -1 & 1 \\ 0 & 3 & 1 \\ 1 & 4 & -2 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}.$$

Agora precisamos saber se o sistema é consistente (isto é, se possui ao menos uma solução) para cada escolha de $\mathbf{u} = (u_1, u_2, u_3)^T$. Se denotarmos a matriz dos coeficientes como A e calcularmos o determinante, temos que $\det(\mathrm{A}) = -24$.

Atenção: deixamos para você conferir!

Como o $\det(\mathrm{A})$ é não nulo, então a matriz dos coeficientes é inversível, e assim o sistema possui sempre solução independentemente da escolha de $\mathbf{u} = (u_1, u_2, u_3)^T$. Isto equivale a dizer que é possível determinar escalares c_1, c_2, c_3, de tal forma que $\mathbf{u}$ (um vetor genérico que representa qualquer vetor de $\mathbb{R}^3$) pode ser escrito como uma combinação linear dos vetores dados $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$, o que nos diz que $\mathbb{R}^3$ está contido no $[\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}]$, e assim temos mostrado que $[\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}] \in \mathbb{R}^3$.

b) Faremos este exemplo rapidamente, pois o procedimento é análogo ao anterior. A matriz correspondente é $\begin{pmatrix} 1 & 3 & -3 \\ 2 & -1 & 8 \\ -1 & 1 & -5 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$, cujo determinante é nulo e, então, é singular. Isso significa que existe, no mínimo, algum $\mathbf{u} = (u_1, u_2, u_3)^T$ para o qual o sistema não terá solução, e então $\mathbf{u} = (u_1, u_2, u_3)^T$ não poderá ser escrito como uma combinação linear destes três vetores.

Observe que, de fato, existirão infinitas escolhas para as quais o sistema não terá solução!

Resumindo, sabemos que $[\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}]$ está contido em $\mathbb{R}^3$, porém, temos mostrado que existe pelo menos um vetor (em verdade infinitos) de $\mathbb{R}^3$ que não está contido em $[\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}]$ e, então, o espaço gerado por estes três vetores não é todo $\mathbb{R}^3$.

Este exemplo nos mostra dois pontos principais.

Em primeiro lugar, nos mostra que nem sempre podemos esperar que quaisquer três vetores de $\mathbb{R}^3$ gerem todo $\mathbb{R}^3$.

Esta idéia será explorada na próxima seção.

O segundo ponto observado é que dois diferentes conjuntos de vetores podem gerar todo $\mathbb{R}^3$.

Temos os três vetores do primeiro item do exemplo anterior, como também o conjunto de vetores canônicos de $\mathbb{R}^3$, gerando $\mathbb{R}^3$. Concluímos que o conjunto de vetores que geram um espaço não é único. Em outras palavras, podemos ter mais de um conjunto de vetores gerando o mesmo espaço.

3.5 Independência Linear

O conceito de geração de um espaço está relacionado com o problema da existência da solução da equação vetorial $v = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n}$, onde $c_1, c_2, \ldots, c_n$ são as incógnitas. Existirá uma única solução $\mathrm{c} = (c_1, c_2, \ldots, c_n)^T$ se, e somente se, a correspondente equação homogênea $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n} = 0$, possuir unicamente a solução trivial, $\mathrm{c} = 0$. Este novo problema conduz à introdução do conceito de independência linear, intimamente relacionado com o conceito de geração de espaços.

Você pode lembrar de uma equivalência idêntica para a solução da equação $A\mathbf{x} = \mathbf{b}$. O que seriam A, $\mathbf{x}$ e $\mathbf{b}$?

Assim, nesta seção vamos olhar mais de perto a estrutura de um espaço vetorial. Para começar, vamos restringir a espaços vetoriais que podem ser gerados por um número finito de elementos. Cada vetor no espaço pode ser construído a partir dos elementos nesse conjunto gerado, usando-se apenas as operações de soma e multiplicação por um escalar. É desejável encontrar um conjunto gerador mínimo, vamos dizer que por economia. Por mínimo, queremos dizer um conjunto gerador sem elementos desnecessários, isto é, para o qual todos os elementos no conjunto são necessários para se gerar o espaço vetorial. Para encontrar um conjunto gerador mínimo, é preciso considerar como os vetores no conjunto dependem um do outro. Esses conceitos simples vão nos dar a chave para entender a estrutura dos espaços vetoriais.

Vamos considerar os seguintes vetores em $\mathbb{R}^3$

$$\mathbf{x_1} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \mathbf{x_2} = \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix}, \mathbf{x_3} = \begin{pmatrix} -1 \\ 3 \\ 8 \end{pmatrix}.$$

Seja S o subespaço gerado por $\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}$. Observe que S pode ser representado, de fato, pelos vetores $\mathbf{x_1}$ e $\mathbf{x_2}$, já que $\mathbf{x_3}$ pertence ao es-

paço gerado por $\mathbf{x_1}$ e $\mathbf{x_2}$, ou seja, $\mathbf{x_3} = 3\mathbf{x_1} + 2\mathbf{x_2}$. Qualquer combinação linear de $\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}$ pode ser reduzida a uma combinação linear de $\mathbf{x_1}$ e $\mathbf{x_2}$:

$$\alpha_1\mathbf{x_1} + \alpha_2\mathbf{x_2} + \alpha_3\mathbf{x_3} = \alpha_1\mathbf{x_1} + \alpha_2\mathbf{x_2} + \alpha_3(3\mathbf{x_1} + 2\mathbf{x_2}) = (\alpha_1 + 3\alpha_3)\mathbf{x_1} + (\alpha_2 + 2\alpha_3)\mathbf{x_2}.$$

Logo, $S = [\{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}\}] = [\{\mathbf{x_1}, \mathbf{x_2}\}]$.

A equação $\mathbf{x_3} = 3\mathbf{x_1} + 2\mathbf{x_2}$ pode ser reescrita na forma $3\mathbf{x_1} + 2\mathbf{x_2} - \mathbf{x_3} = 0$. Como os três coeficientes são diferentes de zero, podemos resolver para qualquer um dos vetores em funções dos outros dois:

$$\begin{aligned} \mathbf{x_1} &= -\frac{2}{3}\mathbf{x_2} + \frac{1}{3}\mathbf{x_3} \\ \mathbf{x_2} &= -\frac{3}{2}\mathbf{x_1} + \frac{1}{2}\mathbf{x_3} \\ \mathbf{x_3} &= 3\mathbf{x_1} + 2\mathbf{x_2}. \end{aligned}$$

Temos, então, que $[\{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}\}] = [\{\mathbf{x_2}, \mathbf{x_3}\}] = [\{\mathbf{x_1}, \mathbf{x_3}\}] = [\{\mathbf{x_1}, \mathbf{x_2}\}]$.

Por causa da relação de dependência, o subespaço S pode ser gerado por qualquer dos dois vetores dados.

Por outro lado, não existe nenhuma relação de dependência entre $\mathbf{x_1}$ e $\mathbf{x_2}$. De fato, se c_1 e c_2 forem escalares, tais que um deles é diferente de 0, e se $c_1\mathbf{x_1} + c_2\mathbf{x_2} = 0$, então poderíamos resolver para um vetor em função do outro como:

$$\mathbf{x_1} = -\frac{c_2}{c_1}\mathbf{x_2}(c_1 \neq 0) \text{ ou } \mathbf{x_2} = -\frac{c_2}{c_1}\mathbf{x_1}(c_2 \neq 0).$$

No entanto, nenhum dos dois vetores em pauta é múltiplo de outro. Logo, $[\{\mathbf{x_1}\}]$ e $[\{\mathbf{x_2}\}]$ são subespaços próprios de S e a equação $c_1\mathbf{x_1} + c_2\mathbf{x_2} = 0$ só será satisfeita se $c_1 = c_2 = 0$.

Podemos generalizar esses exemplos fazendo as seguintes observações:

i) Se $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ gera um espaço vetorial V e um desses vetores pode ser escrito como um combinação linear dos outros $n-1$ vetores, então esses outros $n-1$ vetores geram V.

ii) Dados n vetores $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$, é possível escrever um dos vetores como uma combinação linear de outros $n-1$ vetores se, e somente se, existem escalares $c_1, c_2, \ldots, c_n$ nem todos nulos, tais que $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n} = 0$.

Definição. Os vetores $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$, em um espaço vetorial V, são ditos *linearmente independentes* se a combinação linear dos vetores $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ for igual ao vetor nulo, isto é: $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n} = 0$. Isto implica, necessariamente, que todos os escalares $c_1, c_2, \ldots, c_n$ devem ser iguais a zero.

Como conseqüência de (i) e (ii), vemos que se $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ é um conjunto gerador mínimo, então $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ são linearmente independentes. Em contrapartida, se $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ são linearmente independentes e geram V, então $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ é um conjunto gerador mínimo para V. Um conjunto gerador mínimo é chamado de base.

O conceito de base será estudado com mais detalhes logo adiante.

Exemplo 25. Os vetores $\begin{pmatrix}1\\1\end{pmatrix}$ e $\begin{pmatrix}1\\2\end{pmatrix}$ são linearmente independentes, pois se $c_1\begin{pmatrix}1\\1\end{pmatrix} + c_2\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$, então $\begin{cases} c_1 + c_2 = 0 \\ c_1 + 2c_2 = 0 \end{cases}$. E a única solução desse sistema é $c_1 = c_2 = 0$.

Definição. Os vetores $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$, em um espaço vetorial V, são ditos **linearmente dependentes** se existem escalares $c_1, c_2, \ldots, c_n$ nem todos nulos (pelo menos um deles diferente de zero), tais que $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n} = 0$.

Exemplo 26. Seja $\mathrm{x} = (1,2,3)^T$. Os vetores $\mathbf{x}, \mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}$ são linearmente dependentes, já que a combinação linear nula $\mathbf{e_1} + \mathbf{e_2} + \mathbf{e_3} - \mathbf{x} = 0$ não implica que todos os coeficientes $c_1, c_2, \ldots, c_n$ sejam nulos. (De fato, veja que neste caso $c_1 = 1, c_2 = 2, c_3 = 3, c_4 = -1$).

Definição. Seja $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ um conjunto de no mínimo dois vetores $(n \geq 2)$ em um espaço vetorial V. Então S é linearmente dependente se, e somente se, um desses vetores pode ser escrito como uma combinação linear dos restantes.

As duas últimas definições estabelecem diferentes formas de caracterizar um conjunto linearmente dependente.

Veja, no exemplo anterior, que é imediato escrever $\mathbf{x}$ como uma combinação linear de $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}$.

Dado um conjunto de vetores $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$, em um espaço vetorial V, é trivial encontrar escalares $c_1, c_2, \ldots, c_n$, tais que $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n} = 0$. Basta definir $c_1 = c_2 = \ldots = c_n = 0$.

Se existem escolhas não-triviais de escalares (quer dizer, escalares diferentes de zero) para os quais a combinação linear $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n}$ é igual ao vetor nulo, então o conjunto $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ é *linearmente dependente.*

Se a única maneira de a combinação linear dos vetores $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ ser igual ao vetor nulo for quando todos os escalares $c_1, c_2, \ldots, c_n$ forem iguais a zero, então $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ são *linearmente independentes.*

Uma Interpretação Geométrica da Independência Linear

Um vetor $\mathbf{u}$ é linearmente independente se a combinação linear nula $c\mathbf{u} = 0$ implica necessariamente que $c = 0$. Esta implicação se cumpre se, e somente se, $\mathbf{u} \neq 0$. Então podemos concluir que $\mathbf{u}$ é linearmente independente, ou seja, $\mathbf{u} \neq 0$.

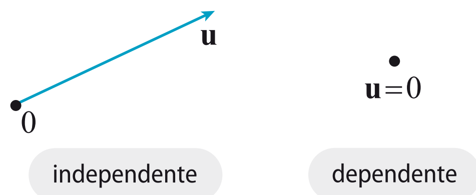


Figura 3.6 – Independente e dependente.

O espaço gerado por um único vetor é uma reta passando pela origem. Span{v} está formado por todos os múltiplos de v, que têm a mesma direção de v. Os pontos extremos desses vetores estão sobre uma reta cuja equação vetorial é $\mathbf{r} = t\mathbf{v} + r_0$, onde t é um escalar.

Dois vetores $\mathbf{u}, \mathbf{v}$ são linearmente *dependentes* se existem α e β, ambos não nulos, tais que $\alpha\mathbf{u} + \beta\mathbf{v} = 0$. Se $\alpha \neq 0$, então temos $\mathbf{u} = -(\beta / \alpha)\mathbf{v}$. Se $\beta \neq 0$, então temos $\mathbf{v} = -(\alpha / \beta)\mathbf{u}$. Assim, podemos concluir que:

Dois vetores são linearmente dependentes se um deles é múltiplo escalar do outro. Em outras palavras, os vetores são paralelos.

Poderia enunciar qual é a equivalência contrária?

Dois vetores são linearmente independentes se os vetores não são paralelos e, então, não são múltiplos entre si.

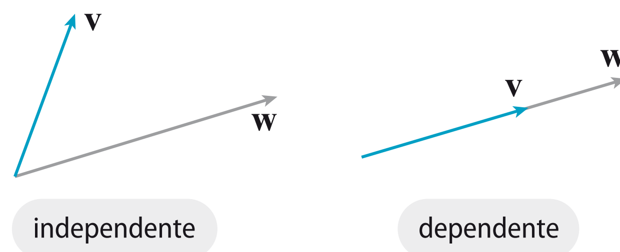


Figura 3.7 – Vetores linearmente independentes e dependentes.

O espaço gerado por dois vetores linearmente independentes é um plano contendo a origem. Para ver isto em $\mathbb{R}^3$, sejam $\mathbf{v}$ e $\mathbf{w}$ dados por $\mathbf{v} = (v_1, v_2, v_3)^T$ e $\mathbf{w} = (w_1, w_2, w_3)^T$, respectivamente. O plano gerado por $\mathbf{v}$ e $\mathbf{w}$ tem um vetor normal dado por $\mathbf{v} \times \mathbf{w}$, cuja equação vetorial é $(\mathrm{v} \times \mathrm{w}) \cdot (x_1 - 0, x_2 - 0, x_3 - 0)^T = 0$, onde $\mathbf{x} = (x_1, x_2, x_3)^T$ é um vetor que pertence ao plano gerado por $\mathbf{v}$ e $\mathbf{w}$. Esta equação, em termos das componentes dos vetores, é dada por:

Isto é, produto vetorial de v e w.

$$(v_2 w_3 - v_3 w_2)x_1 + (v_3 w_1 - v_1 w_3)x_2 + (v_1 w_2 - v_2 w_1)x_3 = 0 \qquad (1)$$

Entretanto, se por outro lado escrevemos $\mathbf{x} = (x_1, x_2, x_3)^T$ como uma combinação linear de $\mathbf{v}$ e $\mathbf{w}, (\mathbf{x} = \alpha v + \beta w)$, temos:

$$\mathbf{x_1} = \alpha v_1 + \beta w_1$$
$$\mathbf{x_2} = \alpha v_2 + \beta w_2$$
$$\mathbf{x_3} = \alpha v_3 + \beta w_3\ .$$

Que pode ser reescrito em forma matricial como, e após escalonamento, como:

$$\begin{pmatrix} 1 & \dfrac{w_1}{v_1} \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \dfrac{x_1}{v_1} \\ \left(v_1x_2 - v_2x_1\right)\left(v_1x_1 - v_3w_1\right) \\ \left(v_1x_3 - v_2x_1\right)\left(v_1w_2 - v_2w_1\right) - \left(v_1x_2 - v_2x_1\right)\left(v_1w_3 - v_3w_1\right) \end{pmatrix}$$

Que possui solução se, e somente se,

$$(v_1x_3 - v_2x_1)(v_1w_2 - v_2w_1) - (v_1x_2 - v_2x_1)(v_1w_3 - v_3w_1) = 0 .$$

Que não é outra coisa que a equação vetorial (1) anterior. Assim, vemos que dois vetores linearmente independentes geram um plano contendo esses vetores.

Para três vetores $\mathbf{u}, \mathbf{v}, \mathbf{w}$ serem linearmente dependentes, α, β, γ, em todos zeros, ficam $\alpha\mathbf{u} + \beta\mathbf{v} + \gamma\mathbf{w} = 0$. Se $\alpha \neq 0$, então $\mathbf{u} = -(\beta/\alpha)\mathbf{v} + -(\gamma/\alpha)\mathbf{w}$ está em um plano, em uma reta ou ainda em um ponto, (dependendo do grau de dependência dos três vetores) gerado por $\mathbf{v}$ e $\mathbf{w}$. Em particular, os três vetores estão no mesmo plano (coplanares). Argumentos similares podem ser colocados nos casos que outros coeficientes sejam zero. Coloque três lápis em uma mesa com os extremos das borrachas juntos para ter um exemplo gráfico de vetores coplanares. Assim:

- Três vetores são linearmente dependentes se estes vetores estão no mesmo plano.

A independência linear de três vetores pode ser entendida pela proposição contrária.

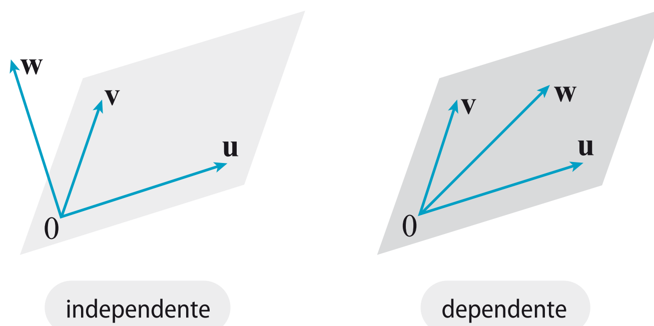


Figura 3.8 – Independência linear.

Estas idéias podem ser estendidas para o caso de vários vetores. As seguintes equivalências podem ser úteis para a compreensão dos conceitos.

- "Vetores linearmente dependentes: significa que um deles é uma combinação dos outros."

Investiguemos a relação entre independência linear e geração de espaços.

- A independência linear de um conjunto de vetores pode ser visualizada geometricamente como um espaço se expandindo na medida em que cada um dos vetores acrescenta uma nova direção, aumentando a dimensão do espaço. Isto pode ser pensado em termos de máxima eficiência na geração do espaço, desde que todos os vetores sejam aproveitados. Ou seja, a eliminação de qualquer vetor produzirá um espaço gerado menor.

3.5.1 Propriedades da Independência Linear

A primeira propriedade está relacionada com a permanência da independência linear quando eliminamos vetores.

Atenção: compare com o efeito na geração de um espaço.

- Se o conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$ é linearmente independente. $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ é linearmente independente.

É fácil entender esta propriedade da independência linear, no sentido que não há desperdício: se uma equipe de cinco pessoas é imprescindível para realizar uma determinada tarefa, então qualquer subgrupo de quatro (ou três, duas ou uma delas) também é necessário, porém, não será capaz de realizar a tarefa em forma completa.

Formulemos a propriedade recíproca da propriedade acima:

- Se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ são linearmente independentes e $\mathbf{v}_{k+1}$ pertence *ao espaço gerado por* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, então $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}$ são linearmente dependentes, de forma similar à geração de espaços.

Os procedimentos mostrados nos exemplos seguintes são úteis para determinar a independência linear de um conjunto de vetores.

Exemplo 27. Neste exemplo, mostraremos que os vetores $\mathbf{v}_1 = (1,-1,0,1)$, $\mathbf{v}_2 = (3,-1,4,3)$, $\mathbf{v}_3 = (2,-1,2,2)$, $\mathbf{v}_4 = (0,1,4,-2)$ e $\mathbf{v}_5 = (1,0,3,0)$ são linearmente dependentes, utilizando operações elementares sobre a matriz:

$$A^T = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \quad \mathbf{v}_4 \quad \mathbf{v}_5] = \begin{pmatrix} 1 & 3 & 2 & 0 & 1 \\ -1 & -1 & -1 & 1 & 0 \\ 0 & 4 & 2 & 4 & 3 \\ 1 & 3 & 2 & -2 & 0 \end{pmatrix}.$$

As colunas de A são os vetores $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \quad \mathbf{v}_4 \quad \mathbf{v}_5]$. Podemos usar as operações elementares por colunas sobre A (ou por linhas *sobre* A^T) para simplificar os vetores e obter a resposta. A forma escalonada por colunas é:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \end{pmatrix}.$$

Observamos agora que os cinco vetores da forma escalonada são claramente linearmente dependentes. Por exemplo, a combinação linear com coeficientes $\mathbf{c}_1 = 0$, $\mathbf{c}_2 = 0$, $\mathbf{c}_3 = 0$, $\mathbf{c}_4 = 0$, $\mathbf{c}_5 = 0$ (nem todos nulos) com as colunas da forma escalonada é igual ao vetor zero, e então o conjunto das colunas é linearmente dependente. Desde que a propriedade de independência linear de um conjunto de vetores não mude pelas operações elementares, concluímos que $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5$ são linearmente dependentes.

O exemplo anterior nos conduz a observações interessantes. Primeiro, para vetores euclidianos, a preservação da independência linear sob operações elementares pode ser revista em termos de operações por colunas.

Se A é equivalente a B por operações por colunas, então colunas de B linearmente independentes correspondem a colunas de A, também linearmente independentes.

Em segundo lugar, observamos uma propriedade do tipo numérico dimensional.

- Se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_\mathbf{n} \in \mathbb{R}^n$ são linearmente independentes, então $k \leq n$.

Aproveitemos este exemplo para introduzir o seguinte teorema. Se considerarmos apenas os vetores $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4}$, teremos:

$$\mathrm{A} = [\mathbf{v_1} \quad \mathbf{v_2} \quad \mathbf{v_3} \quad \mathbf{v_4}] = \begin{pmatrix} 1 & 3 & 2 & 0 \\ -1 & -1 & -1 & 1 \\ 0 & 4 & 2 & 4 \\ 1 & 3 & 2 & -2 \end{pmatrix}.$$

E agora a sua respectiva forma escalonada por colunas é:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 1 & 0 & -1 & 0 \end{pmatrix}.$$

Claramente esta matriz é singular.

Teorema. *Sejam $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}$, n vetores em $\mathbb{R}^n$, com $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ para $i = 1, \ldots, n$. Se $\mathrm{A} = [\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}]$, então os vetores $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}$ são linearmente dependentes se, e somente se, A é singular.*

Demonstração. A combinação linear nula $c_1\mathrm{x}_1 + c_2\mathrm{x}_2 + \ldots + c_n\mathrm{x}_n = 0$ é equivalente ao sistema de equações:

$$\begin{aligned} c_1\mathrm{x}_{11} + c_2\mathrm{x}_{12} + \ldots + c_n\mathrm{x}_{1n} &= 0 \\ c_1\mathrm{x}_{21} + c_2\mathrm{x}_{22} + \ldots + c_n\mathrm{x}_{2n} &= 0 \\ &\vdots \\ c_1\mathrm{x}_{n1} + c_2\mathrm{x}_{n2} + \ldots + c_n\mathrm{x}_{nn} &= 0. \end{aligned}$$

Definido $\mathbf{c} = (c_1, c_2, ..., c_n)^T$, o sistema pode ser escrito em forma matricial como $A\mathrm{c} = 0$.

Essa equação tem uma solução não trivial se, e somente se, a A é singular. Portanto, $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}$ são linearmente dependentes se, e somente se, A é singular.

Podemos usar o teorema anterior para testar se n vetores são linearmente independentes em $\mathbb{R}^n$. Basta formar a matriz A, cujas colunas são os vetores a serem testados. Para determinar se A é ou não singular, basta calcular o valor do determinante de A. Se $\det(\mathrm{A}) = 0$, os vetores são linearmente dependentes. Se $\det(\mathrm{A}) \neq 0$, os vetores são linearmente independentes.

Voltando ao exemplo anterior, temos que:

$$\det(A) = \begin{vmatrix} 1 & 3 & 2 & 0 \\ -1 & -1 & -1 & 1 \\ 0 & 4 & 2 & 4 \\ 1 & 3 & 2 & -2 \end{vmatrix} = 0$$

e então os vetores são linearmente dependentes.

Na prática, para determinar se um conjunto de vetores é ou não linearmente independente em $\mathbb{R}^n$, precisamos resolver um sistema homogêneo de equações lineares.

Exemplo 28. Determine se $S = \{1+x, x+x^2, 1+x^2\}$ é linearmente independente em P_2.

Solução. Considere a seguinte combinação linear:

$$\alpha(1+x)+\beta(x+x^2)+\gamma(1+x^2)=0.$$

Reagrupando convenientemente os termos do lado esquerdo da equação, podemos escrever

$$(\alpha+\gamma)+(\alpha+\beta)x+(\beta+\gamma)x^2 = 0$$
$$0\cdot 1+0\cdot x+0\cdot x^2 = 0.$$

Sabemos que um polinômio é identicamente nulo quando todos os seus coeficientes são zero. Assim, igualando os respectivos coeficientes das potências de x, temos: $\alpha+\gamma=0$, $\alpha+\beta=0$, $\beta+\gamma=0$, que possui apenas a solução trivial $\alpha=\beta=\gamma=0$.

Logo, S é linearmente independente.

Exemplo 29. Mostre que o conjunto de $M_{2\times 2}$, formado por $S = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ 6 & 9 \end{pmatrix} \right\}$, é um conjunto linearmente dependente e, em seguida, escreva um destes vetores como uma combinação linear dos outros.

Solução. Considere a seguinte cominação linear nula

$$\alpha\begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}+\beta\begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix}+\gamma\begin{pmatrix} -1 & 2 \\ 6 & 9 \end{pmatrix}=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Esta equação é equivalente ao sistema:

$$\begin{aligned} \alpha-\beta-\gamma&=0\\ \alpha-2\gamma&=0\\ 2\beta+6\gamma&=0\\ 3\alpha+\beta+9\gamma&=0\,, \end{aligned}$$

Atenção: verifique essa solução!

onde $\alpha=-2k$, $\beta=-3k$, $\gamma=k$ é solução, sendo k arbitrário.

Logo, S é linearmente dependente. Por exemplo, se escolhermos $k=1$, temos que:

$$-2\begin{pmatrix}1&1\\0&3\end{pmatrix}-3\begin{pmatrix}-1&0\\2&1\end{pmatrix}+\begin{pmatrix}-1&2\\6&9\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix},$$

e assim, podemos escrever

$$\begin{pmatrix}1&1\\0&3\end{pmatrix}=-\frac{3}{2}\begin{pmatrix}-1&0\\2&1\end{pmatrix}+\frac{1}{2}\begin{pmatrix}-1&2\\6&9\end{pmatrix},$$

ou também

$$\begin{pmatrix}-1&2\\6&9\end{pmatrix}=2\begin{pmatrix}1&1\\0&3\end{pmatrix}+3\begin{pmatrix}-1&0\\2&1\end{pmatrix},$$

ou ainda

$$\begin{pmatrix}-1&0\\2&1\end{pmatrix}=-\frac{2}{3}\begin{pmatrix}1&1\\0&3\end{pmatrix}+\frac{1}{3}\begin{pmatrix}-1&2\\6&9\end{pmatrix}.$$

Exercícios Propostos

1) Determine se os vetores dados são ou não linearmente independentes em $\mathbb{R}^2$.

a) $\begin{pmatrix}2\\1\end{pmatrix},\begin{pmatrix}3\\2\end{pmatrix}$

b) $\begin{pmatrix}2\\3\end{pmatrix},\begin{pmatrix}4\\6\end{pmatrix}$

c) $\begin{pmatrix}-2\\1\end{pmatrix},\begin{pmatrix}1\\3\end{pmatrix},\begin{pmatrix}2\\4\end{pmatrix}$

d) $\begin{pmatrix}-1\\2\end{pmatrix},\begin{pmatrix}1\\-2\end{pmatrix},\begin{pmatrix}2\\-4\end{pmatrix}$

e) $\begin{pmatrix}1\\2\end{pmatrix},\begin{pmatrix}-1\\1\end{pmatrix}$

2) Determine se os vetores são ou não linearmente independentes em $\mathbb{R}^3$.

a) $\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\1\end{pmatrix},\begin{pmatrix}1\\0\\1\end{pmatrix}$

d) $\begin{pmatrix}2\\1\\-2\end{pmatrix},\begin{pmatrix}-2\\-1\\2\end{pmatrix},\begin{pmatrix}4\\2\\-4\end{pmatrix}$

b) $\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\1\end{pmatrix},\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}1\\2\\3\end{pmatrix}$

c) $\begin{pmatrix}2\\1\\-2\end{pmatrix},\begin{pmatrix}3\\2\\-2\end{pmatrix},\begin{pmatrix}4\\2\\-4\end{pmatrix}$

e) $\begin{pmatrix}1\\1\\3\end{pmatrix},\begin{pmatrix}0\\2\\1\end{pmatrix}$

3) Descreva geometricamente o espaço gerado por cada um dos seguintes vetores no exercício 2.

4) Determine se os vetores dados são ou não linearmente independente em $\mathbb{R}^{2\times 2}$.

a) $\begin{pmatrix}1&0\\1&1\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix}$

b) $\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}$

c) $\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}2&3\\0&2\end{pmatrix}$

5) Determine se os vetores dados são ou não linearmente independentes em P_3.

a) $1, x^2, x^2-2$

b) $2, x^2, x, 2x+3$

c) $x+2, x+1, x^2-1$

d) $x+2, x^2-1$

3.6 Bases e Dimensão

O conjunto de vetores $S=\{(1,1)^T,(1,-1)^T\}$ gera $\mathbb{R}^2$, isto é, qualquer vetor em $\mathbb{R}^2$ pode ser obtido como uma combinação linear de $(1,1)^T$ e $(1,-1)^T$.

O conjunto de vetores $T=\{(1,1)^T,(1,-1)^T,(1,0)^T\}$ também gera $\mathbb{R}^2$. Os conjuntos S e T diferem entre si: S é linearmente independente, enquanto T é linearmente dependente, o que faz diferença quando escrevemos um vetor como uma combinação linear dos vetores de cada um dos conjuntos.

Por exemplo, para escrevermos $(2,4)^T$ em termos dos vetores de S, conseguimos de uma única maneira: $(2,4)^T=3(1,1)^T-(1,-1)^T$.

Entretanto, em termos dos vetores de T, temos várias possibilidades:

$$(2,4)^T = 3(1,1)^T - (1,-1)^T + 0(1,0)^T$$
$$(2,4)^T = 0(1,1)^T - 4(1,-1)^T + 6(1,0)^T$$
$$(2,4)^T = 4(1,1)^T + 0(1,-1)^T - 2(1,0)^T$$

ou, em geral, $(2,4)^T = (k+4)(1,1)^T + k(1,-1)^T + (-2-2k)(1,0)^T$.

O ponto chave é: se um conjunto de vetores gera V e esse conjunto for linearmente dependente, então a representação de um vetor x em termos dos vetores desse conjunto não é única.

Para se ter *unicidade*, o conjunto gerador deve ser também linearmente independente. Um conjunto assim é chamado de uma base para V.

Em geral, considere a equação vetorial $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n} = v$ em um espaço vetorial V, onde $c_1, c_2, \ldots, c_n$ são as incógnitas. A existência de soluções para todo $\mathbf{v} \in V$ é equivalente a ter V *gerado pelos vetores* $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$. Entretanto, a propriedade de independência linear sobre $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ garantirá a unicidade da solução.

Definição. Os vetores $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ formam uma base para o espaço vetorial V se, e somente se, $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ são linearmente independentes e geram V.

Assim, para determinar se um conjunto de vetores é uma base para V, é suficiente verificar se o conjunto gera V e é linearmente independente.

Exemplo 30. Mostre que o conjunto $S = \{(1,2)^T, (3,-1)^T\}$ é uma base para $\mathbb{R}^2$.

Solução. Devemos mostrar que S é linearmente independente, o que equivale a mostrar que $\alpha(1,2)^T + \beta(3,-1)^T = (0,0)^T$ tem como única solução $\alpha = \beta = 0$, ou que $\alpha(1,2)^T + \beta(3,-1)^T = (a,b)^T$ tem uma única solução para qualquer $(a,b)^T$.

Estas equações podem ser escritas matricialmente (em sua forma aumentada) como:

$$\begin{pmatrix} 1 & 3 & \vdots & a \\ 2 & -1 & \vdots & b \end{pmatrix} \text{ e } \begin{pmatrix} 1 & 3 & \vdots & 0 \\ 2 & -1 & \vdots & 0 \end{pmatrix}.$$

Em vez de resolver ambos os sistemas separadamente, (por economia) o resolvemos simultaneamente, trabalhando com uma matriz duplamente aumentada:

$$\begin{pmatrix} 1 & 3 & \vdots & a & \vdots & 0 \\ 2 & -1 & \vdots & b & \vdots & 0 \end{pmatrix}.$$

Escalonando, obtemos:

$$\left(\begin{array}{c} \overbrace{\begin{matrix} 1 & 3 & \vdots & a & \vdots \\ 0 & 1 & \vdots & (-b+2a)/7 & \vdots \end{matrix}}^{\text{gereção}} \begin{matrix} 0 \\ 0 \end{matrix} \end{array}\right).$$

(com a chave inferior "independência linear" abrangendo as colunas $1, 3$ e 0)

Achamos, para geração, que $\alpha=(a+3b)/7, \beta=(2a-b)/7$ e, para independência linear, que $\alpha=\beta=0$. Já que S é linearmente independente e gera $\mathbb{R}^2$, é uma base para $\mathbb{R}^2$.

Neste exemplo, os coeficientes da combinação linear são únicos para um vetor $(a,b)^T$ dado. O que é verdade, em geral.

Teorema. *Seja* $S=\{\mathbf{v_1},\mathbf{v_2},\ldots,\mathbf{v_n}\}$ *uma base para um espaço vetorial* V. *Seja* v *um vetor em* V. *Os coeficientes na representação* $v=c_1\mathbf{v_1}+c_2\mathbf{v_2}+\ldots+c_n\mathbf{v_n}$ *são únicos.*

Prova. Suponhamos que temos duas representações para $\mathbf{v}$: $\mathbf{v}=a_1\mathbf{v_1}+a_2\mathbf{v_2}+\ldots+a_n\mathbf{v_n}$ e $\mathbf{v}=b_1\mathbf{v_1}+b_2\mathbf{v_2}+\ldots+b_n\mathbf{v_n}$. Mostraremos que os coeficientes são iguais.

Para isto, formamos a soma $\mathbf{v}+(-\mathbf{v})$, que é igual a 0. Recombinando os diferentes termos, obtemos

$$0=\mathbf{v}+(-\mathbf{v})=(a_1-b_1)\mathbf{v_1}+(a_2-b_2)\mathbf{v_2}+\ldots+(a_n-b_n)\mathbf{v_n}.$$

Desde que S seja uma base com um conjunto de vetores linearmente independentes, os coeficientes da combinação linear devem ser todos iguais a zero. Isto é, $a_1=b_1, a_2=b_2,\ldots,a_n=b_n$, os coeficientes da combinação linear original são os mesmos.

Exemplo 31. Os vetores canônicos $\mathbf{e_1}=(1,0,0)$, $\mathbf{e_2}=(0,1,0)$ e $\mathbf{e_3}=(0,0,1)$ geram $\mathbb{R}^3$ e são linearmente independentes. Logo, o conjunto $\{\mathbf{e_1},\mathbf{e_2},\mathbf{e_3}\}$ é uma base para $\mathbb{R}^3$. Este fato se estende naturalmente para os vetores canônicos de $\mathbb{R}^n$.

Exemplo 32. Os monômios $1,t,\ldots,t^n$ geram P_n e são linearmente independentes. Logo, são uma base para P_n.

Exemplo 33. As seguintes igualdades:

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}=a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}+b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}+c\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}+d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}=a\mathrm{E}_{11}+b\mathrm{E}_{12}+c\mathrm{E}_{21}+d\mathrm{E}_{22}$$

têm duas implicações. Qualquer matriz 2×2 é uma combinação linear de $\mathrm{E}_{11},\mathrm{E}_{12},\mathrm{E}_{21},\mathrm{E}_{22}$. Se a combinação linear de $\mathrm{E}_{11},\mathrm{E}_{12},\mathrm{E}_{21},\mathrm{E}_{22}$ (com a,b,c,d como coeficientes) é a matriz nula, então todos os coeficientes de A (que são a,b,c,d) devem ser zero.

A primeira implicação significa exatamente que $M_{2\times 2}$ é gerado por $\mathrm{E}_{11},\mathrm{E}_{12},\mathrm{E}_{21},\mathrm{E}_{22}$ e a segunda que $\mathrm{E}_{11},\mathrm{E}_{12},\mathrm{E}_{21},\mathrm{E}_{22}$ são linearmente independentes. Logo, $\{\mathrm{E}_{11},\mathrm{E}_{12},\mathrm{E}_{21},\mathrm{E}_{22}\}$ é uma base para $M_{2\times 2}$.

Em geral, seja E_{ij} uma matriz $m\times n$, cujas entradas são todas iguais a zero, exceto na entrada correspondente a i,j, que é igual a um para qualquer vetor $\mathbf{b}=(b_1,b_1,b_1)^T$. Então, $\{\mathrm{E}_{ij}:1\le i\le m, 1\le j\le n\}$ é uma base para $M_{m\times n}$.

Exemplo 34. O sistema

$$\begin{aligned} 3x_1+x_2-x_3&=2\\ x_1-x_2+x_3&=2\\ 2x_1+2x_2+x_3&=6 \end{aligned}$$

tem uma única solução (confira!). Já que o número de linhas e colunas é igual, sabemos que o sistema mais geral:

$$\begin{aligned} 3x_1+x_2-x_3&=b_1\\ x_1-x_2+x_3&=b_2\\ 2x_1+2x_2+x_3&=b_3 \end{aligned}$$

possui uma única solução para qualquer vetor $\mathbf{b}=(b_1,b_1,b_1)^T$ (por quê?). Já que esse sistema é a mesma coisa que a equação vetorial:

$$\mathbf{x_1}(3,1,2)+\mathbf{x_2}(1,-1,2)+\mathbf{x_3}(-1,1,1)=(b_1,b_2,b_3)$$

e isto implica que os vetores $(3,1,2)$, $(1,-1,2)$, $(-1,1,1)$ geram $\mathbb{R}^3$, e a unicidade implica que os três vetores são linearmente independentes. Concluímos, então, que os três vetores formam uma base para $\mathbb{R}^3$.

Os argumentos utilizados neste último exemplo se aplicam, em geral, a conjuntos de vetores euclidianos. Uma base para $\mathbb{R}^n$ sempre significa a existência de n vetores linearmente independentes.

Se as colunas de uma matriz A formam uma base para $\mathbb{R}^n$, então A deve ser uma matriz $n \times n$. Você precisa de n vetores de $\mathbb{R}^n$ linearmente independentes para gerar todo $\mathbb{R}^n$. É interessante a seguinte equivalência:

- Colunas de A formam uma base de $\mathbb{R}^n \Leftrightarrow A\mathbf{x} = \mathbf{b}$ tem uma única solução para qualquer b.

Exemplo 35. Em muitas aplicações, é necessário encontrar um subespaço particular de um espaço vetorial. Isto pode ser feito encontrando-se os elementos de uma base para o subespaço. Por exemplo, para encontrar todas as soluções do sistema

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ 2x_1 + x_2 + x_4 &= 0 \end{aligned}$$

precisamos encontrar o núcleo da matriz $A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix}$, que é o subespaço de $\mathbb{R}^4$ gerado pelos vetores:

$$\begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} \text{ e } \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

Como esses dois vetores são linearmente independentes, eles formam uma base para $\mathbb{R}^4$.

Exemplo 36. Mostre que o conjunto $S = \{(1,2)^T, (3,-1)^T, (1,0)^T\}$ não é uma base para $\mathbb{R}^2$.

> **Solução.** O conjunto S é linearmente dependente, já que, por exemplo, $(1,2)^T + 2(3,-1)^T - 7(1,0)^T = (0,0)^T$. Então S não pode ser uma base para $\mathbb{R}^2$.

Comentário. Apenas temos que mostrar que uma das condições dadas na definição de base não é satisfeita, neste caso a independência linear.

Exemplo 37. O espaço vetorial formado apenas pelo vetor zero não possui uma base porque qualquer subconjunto de vetores, incluindo o vetor zero, é linearmente dependente.

Para mostrar que nem todo espaço possui uma base, precisamos decidir sob quais condições um espaço vetorial dado possui uma base ou não. O conceito de conjuntos de vetores geradores nos ajudará a responder essa questão.

Teorema. *Se $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ é um conjunto de vetores não nulos que geram um subespaço W de um espaço vetorial V, então algum subconjunto de S é uma base para W (que pode ser o próprio S).*

> **Prova.** Se S for um conjunto linearmente independente, então, por definição, S é uma base para W. Se S for linearmente dependente, então, no mínimo, um dos vetores pode ser escrito como uma combinação linear dos outros. Suponhamos que $\mathbf{v_n}$ seja tal vetor (se não for o caso, role os vetores em S até que isto seja verdade). Postulamos que $S' = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_{n-1}}\}$ continua gerando W.
>
> Para ver isto, seja um vetor $\mathbf{x}$ em W que escrevemos como:
>
> $$\mathbf{x} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_{n-1}\mathbf{v_{n-1}} + c_n\mathbf{v_n}\,.$$
>
> Agora, como $\mathbf{v_n} = d_1\mathbf{v_1} + d_2\mathbf{v_2} + \ldots + d_{n-1}\mathbf{v_{n-1}}$ (é o vetor que estamos supondo ser uma combinação linear dos outros), substituindo na expressão anterior e remanejando a equação, temos:
>
> $$\mathbf{x} = (c_1 - c_n d_1)\mathbf{v_1} + (c_2 - c_n d_2)\mathbf{v_2} + \ldots + (c_{n-1} - c_n d_{n-1})\,.$$
>
> Logo, S' gera W.
>
> Se S' é linearmente independente, S' é uma base de W. Se S' é linearmente dependente, um dos vetores de S' é uma combinação linear dos outros.

Agora, podemos argumentar como antes e repetir o processo, eliminando vetores até encontrar fatalmente um conjunto linearmente independente que gere W. (Se reduzirmos o conjunto a um único vetor, esse conjunto é linearmente independente, já que S foi definido como um conjunto de vetores não nulos). O conjunto resultante é uma base para W.

Então, temos um resultado fundamental:

Qualquer espaço vetorial finito gerado por um conjunto de vetores não nulos possui uma base.

Exemplo 38. Seja V o conjunto de todos os polinômios munido com as operações usuais. O espaço vetorial gerado V não é finito. De fato, se tomamos qualquer subconjunto finito S de V, então haverá um termo de grau máximo, digamos t^p, no conjunto. O polinômio t^{p+1} não está agora no espaço gerado S, e então S não pode gerar V.

Restringiremos o nosso estudo às bases de espaços vetoriais finitos.

No Exemplo 36, mostramos que qualquer conjunto de três ou mais vetores de $\mathbb{R}^2$ não pode ser uma base para $\mathbb{R}^2$. Esses vetores poderiam ser coplanares e formar um conjunto linearmente dependente. Um conjunto de apenas um vetor (três é demais, um é pouco) tampouco pode ser uma base para $\mathbb{R}^2$, já que gera apenas uma reta através da origem. Podemos concluir, então, que qualquer base para $\mathbb{R}^2$ deve conter exatamente dois vetores. Veja o seguinte teorema.

Teorema. *Se $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ é uma base para V, então qualquer conjunto de $n+1$ (ou mais) vetores é linearmente dependente, e então, não é uma base para V. Por outro lado, qualquer conjunto de $n-1$ (ou menos) vetores não é suficiente para gerar V, e então, não é uma base para V.*

Prova. i) Seja $T = \{c_1\mathbf{w_1} + c_2\mathbf{w_2} + \ldots + c_n\mathbf{w_n} + c_{n+1}\mathbf{w_{n+1}}\}$, isto é, T contém exatamente $n+1$ vetores. Mostraremos que T não pode ser uma base, evidenciando que T é linearmente dependente. Para isso, consideremos

$$c_1\mathbf{w_1} + c_2\mathbf{w_2} + \ldots + c_n\mathbf{w_n} + c_{n+1}\mathbf{w_{n+1}} = 0 \qquad (1)$$

Cada $\mathbf{w}_k$ pode ser escrito como

$$\mathbf{w_k} = a_{1k}\mathbf{v_1} + a_{2k}\mathbf{v_2} + \ldots + a_{nk}\mathbf{v_n},\ k = 1, 2, \ldots, n+1,$$

desde que S gera V. Substituindo cada w_k na equação (1), temos:

$$\sum_{k=1}^{n+1} c_k\mathbf{w_k} = \sum_{k=1}^{n+1} c_k \sum_{j=1}^{n} a_{jk}\mathbf{v_j} = \sum_{j=1}^{n} \left(\sum_{k=1}^{n+1} a_{jk}c_k\right)\mathbf{v_j} = 0.$$

Já que S é linearmente independente.

$$\sum_{k=1}^{n+1} a_{jk} c_k = 0, \ \forall j = 1, \ldots, n$$

Este é um sistema de equações homogêneo que tem menos equações (n) que incógnitas $(n+1)$; de tal forma, existe uma solução distinta da trivial para $c_1, c_2, \ldots, c_n, c_{n+1}$. Isto significa que T é linearmente dependente.

Agora, suponhamos que T contenha mais que $(n+1)$ vetores. Seja $\widetilde{T}$ um subconjunto de $(n+1)$ vetores, que deve ser linearmente dependente (como acabamos de mostrar). Como $\tilde{T} \subset T$, o conjunto T contém um subconjunto linearmente dependente, e então deve ser linearmente dependente.

ii) Suponhamos agora que T contém $n-1$ vetores e gera V.

Então, por um teorema anterior, T deve possuir uma base τ para V. Se τ contém r vetores, devemos ter que $r \leq n-1$. Já que τ é uma base e S tem $r+1$ ou mais vetores, devemos concluir, levando em consideração o item i, que S é linearmente dependente. Isto contradiz o fato que S é uma base.

Neste ponto, sabemos que um espaço vetorial finito gerado por um conjunto de vetores não nulos tem uma base e é de dimensão finita.

O teorema diz que o número de vetores em uma base é único. Se acharmos uma base S para V e S tiver sete vetores, então qualquer base deverá ter sete vetores. Podemos ter infinitas bases para V, porém cada uma delas terá apenas sete vetores. Este fato nos leva a definir a *dimensão* de um espaço vetorial como sendo o número de vetores que formam a base para V.

Definição. Se uma base S tem n vetores, a dimensão de V é n, e escrevemos $\dim(V) = n$, e dizemos V é de *dimensão finita*. Em particular, V é chamado como um *espaço vetorial n-dimensional* quando a base para V tem n vetores.

A dimensão do espaço vetorial contendo o vetor zero, unicamente, está definida como sendo zero. Agora vamos aos exemplos. Estaremos interessados em determinar uma base e sua dimensão para diferentes espaços vetoriais. Veja que o procedimento será repetido

em cada caso: propor uma base (a partir de uma heurística) e testar geração e independência linear. Nos seguintes exemplos, mostram-se as bases canônicas de diferentes espaços vetoriais.

Exemplo 39. Mostre que $\mathbb{R}^n$ tem a base (canônica) $\{\mathbf{e_1}, \mathbf{e_2}, ..., \mathbf{e_n}\}$, onde

$$\mathbf{e_j} = \left(0, \ldots, \underset{\substack{j-esima \\ componente}}{1}, \ldots, 0\right)^T .$$

Solução. Para testar geração, consideremos qualquer vetor $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ e observemos que $\mathbf{x} = x_1\mathbf{e_1} + x_2\mathbf{e_2} + \ldots + x_n\mathbf{e_n}$.

Para testar independência linear, consideremos a combinação linear nula $c_1\mathbf{e_1} + c_2\mathbf{e_2} + \ldots + c_n\mathbf{e_n} = 0$. Logo, encontramos que $c_1 = c_2 = \ldots = c_n = 0$. Assim, $\{\mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_n}\}$ é uma base para $\mathbb{R}^n$, e $\dim(\mathbb{R}^n) = n$.

É natural associar a base canônica com alguns espaços que nos são familiares:

- $\mathbb{R}^1 \to$ um objeto unidimensional
- $\mathbb{R}^2 \to$ um objeto bidimensional
- $\mathbb{R}^3 \to$ um objeto tridimensional

Comentário. Observando alguns dos exemplos anteriores, vemos que existem diversas bases possíveis para $\mathbb{R}^2$. Em geral, para um espaço vetorial finito, não nulo, existe um número infinito de bases. Entretanto, o número de elementos, em qualquer uma dessas bases, é sempre o mesmo, e lembre-se que este número é a dimensão do espaço.

Exemplo 40. Mostre que $M_{2\times3}$ tem dimensão 6.

Solução. A base canônica é

$$S = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}.$$

Assim, $\dim\left(M_{2\times3}\right) = 6$, igual ao número de vetores em S.

Exemplo 41. O espaço vetorial P_n, polinômios de grau n, tem dimensão $n+1$.

Solução. Uma base é $S = \{1, x, \ldots, x^n\}$. Confiramos primeiro a independência linear. A equação $a_0 1 + a_1 x + \ldots + a_{n-1} x^n = 0$ se satisfaz somente se o polinômio do lado esquerdo for zero para todo número real x. E isto acontece somente se todos os coeficientes são nulos, isto é, somente se $a_1 = a_2 = \ldots = a_n = 0$. Logo, S é linearmente independente. Pode ser visto que S gera P_n pelo fato de que qualquer polinômio em P_n é da forma $a_0 1 + a_1 x + \ldots + a_{n-1} x^n$.

Agora que sabemos o que é base de um espaço vetorial, iremos colocar o que se conhece como *o segundo problema fundamental da álgebra linear.*

O Problema das Bases

Seja V um espaço vetorial. O problema das bases (PB) pode ser colocado de duas maneiras:

Problema 1. Construir uma base para V, selecionando vetores de V.

Problema 2. Dado um conjunto S de vetores em V, construir uma base para V acrescentando ou eliminando alguns (mas não todos) vetores de S, ou algumas vezes ambas as coisas.

Antes de começar a resolver estes problemas, poderíamos nos perguntar se há alguma garantia de que realmente exista resposta para eles. Lembre-se do Teorema da página 172, que nos diz: "jogue fora os vetores dependentes para obter um conjunto gerador e assim obter uma base", o qual poderá nos ajudar nesse momento.

Exemplo 42. *(Problema 2 do PB).* Seja $S = \{(1,0,3)^T, (2,1,4)^T\}$. Ache uma base T para $\mathbb{R}^3$ que contenha S.

Solução 1. Já que $\mathbb{R}^3$ tem dimensão 3, sabemos que T deve conter exatamente três vetores. O conjunto S já é linearmente independente, assim, devemos acrescentar somente um vetor a este conjunto. O novo vetor que juntaremos aos do conjunto S deve ser tal que o conjunto T seja linearmente independente. Isto significa que o novo vetor não deve pertencer ao espaço gerado pelos vetores que já estavam em S.

Então, primeiro determinamos o espaço gerado pelos vetores:

Deixamos pra você conferir!

$$[(1,0,3)^T, (2,1,4)^T] = \{(\alpha + 2\beta, \beta, 3\alpha + 4\beta)^T\}.$$

E agora o que devemos garantir é que o novo vetor não seja da forma $(\alpha+2\beta,\beta,3\alpha+4\beta)^T$.

Vejamos como fazer isto. Suponhamos que o nosso novo vetor seja $\mathbf{x}=(x_1,x_2,x_3)^T$, e forçamos a equação $(\alpha+2\beta,\beta,3\alpha+4\beta)^T=(x_1,x_2,x_3)^T$ para não ter solução para α e β. Isto nos leva, resolvendo esse sistema, a escolher $\mathbf{x}$, de tal forma que $x_3-3x_1+2x_2\neq 0$, para conseguirmos ter um vetor que não pertença $[(1,0,3)^T,(2,1,4)^T]$. Logo, $(0,1,0)^T$ funciona (de fato, existem infinitas escolhas), e então o conjunto $\{(1,0,3)^T,(2,1,4)^T,(0,1,0)^T\}$ é uma base para $\mathbb{R}^3$.

Solução 2. Se o terceiro vetor $\mathbf{x}=(x_1,x_2,x_3)^T$ fosse escolhido de tal forma que $\{(1,0,3)^T,(2,1,4)^T,(x_1,x_2,x_3)^T\}$ seja linearmente dependente, então:

$$\begin{vmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ x_1 & x_2 & x_3 \end{vmatrix}=0$$

devido a uma das linhas ser uma combinação linear das outras. Como o nosso propósito é justamente o contrário, se requerermos que

$$\begin{vmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ x_1 & x_2 & x_3 \end{vmatrix}\neq 0,$$

obtemos novamente $x_3-3x_1+2x_2\neq 0$, que é a mesma condição obtida na Solução 1.

Solução 3. *(Tentativa e erro)*

Chutaremos como candidato ao terceiro vetor os vetores canônicos, até que algum deles funcione. Tentaremos inicialmente com $(1,0,0)^T$ e conferiremos a independência linear. Para isso, formamos a seguinte combinação linear nula:

Atenção: Tente você com outro vetor canônico!

$$\alpha(1,0,3)^T+\beta(2,1,4)^T+\gamma(1,0,0)^T=(0,0,0)^T,$$

o que nos leva a resolver

$$\begin{pmatrix} 1 & 2 & 1 & \vdots & 0 \\ 0 & 1 & 0 & \vdots & 0 \\ 3 & 4 & 0 & \vdots & 0 \end{pmatrix}\rightarrow\begin{pmatrix} 1 & 2 & 1 & \vdots & 0 \\ 0 & 1 & 0 & \vdots & 0 \\ 0 & 0 & 1 & \vdots & 0 \end{pmatrix}.$$

Nossa primeira tentativa teve sucesso, os vetores são linearmente independentes e formam a base $\{(1,0,3)^T, (2,1,4)^T, (1,0,0)^T\}$.

Neste exemplo usamos implicitamente o seguinte teorema.

Teorema: Seja $\dim(V) = n$ e $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_\mathbf{n}\}$ um subconjunto de V. As seguintes afirmativas são equivalentes:

1) O conjunto S é uma base para V,

2) O conjunto S é linearmente independente,

3) O conjunto S gera V.

Observe que, se tivéssemos disposto deste teorema, os exemplos anteriores teriam sido resolvidos com a metade do trabalho. Seria suficiente mostrar independência linear.

Exemplo 43. *(Problema 1 do PB).* Ache uma base para o espaço solução de:

$$\begin{aligned} x_1 + x_2 - x_3 + 2x_4 &= 0 \\ x_2 + x_3 - x_4 &= 0 \\ 3x_1 + 4x_2 - 2x_3 + 5x_4 &= 0\,. \end{aligned}$$

Solução: Resolvendo o sistema, temos que $x_3 = \alpha$, $x_4 = \beta$, e então

Deixamos pra você conferir!

$$(x_1, x_2, x_3, x_4)^T = \alpha(2,-1,0,1)^T + \beta(-3,1,1,0)^T.$$

Já que $S = \{(2,-1,0,1)^T + (-3,1,1,0)^T\}$ gera o espaço solução e é linearmente independente, S é uma base para o espaço solução.

Estes dois últimos exemplos mostram que resolver o PB não segue um procedimento padrão ou a utilização de uma fórmula. Requer habilidade, versatilidade e uma familiaridade com vários conceitos anteriores.

Exemplo 44. O conjunto

$$\text{span}(S) = [(1,-1,2)^T, (0,5,-8)^T, (3,2,-2)^T, (8,2,0)^T]$$

é um espaço vetorial. Ache uma base para esse espaço.

Solução. Devemos eliminar os vetores que são combinação linear dos outros. Para descobrir essa dependência, consideremos a seguinte combinação linear:

$$a(1,-1,2)^T + b(0,5,-8)^T + c(3,2,-2)^T + d(8,2,0)^T = (0,0,0)^T \qquad (1)$$

Se escrevemos esse sistema em forma matricial e escalonamos por linhas, temos:

$$\begin{pmatrix} 1 & 0 & 3 & 8 & \vdots & 0 \\ -1 & 5 & 2 & 2 & \vdots & 0 \\ 2 & -8 & -2 & 0 & \vdots & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 3 & 8 & \vdots & 0 \\ 0 & 1 & 1 & 2 & \vdots & 0 \\ 0 & 0 & 0 & 0 & \vdots & 0 \end{pmatrix}.$$

As soluções são $d = d, c = c, b = -c - 2d, a = -3c - 8d$. Temos duas variáveis livres.

Escolhendo $d = 1$, $c = 0$, temos $(8,2,0)^T = 8(1,-1,2)^T + 2(0,5,-8)^T$.
Escolhendo $d = 0$, $c = 1$, temos $(3,2,-2)^T = 3(1,-1,2)^T + (0,5,-8)^T$.
Logo, $(8,2,0)^T$ e $(3,2,-2)^T$ dependem de $(1,-1,2)^T$ e $(0,5,-8)^T$, e então o $\text{span}(S) = [(1,-1,2)^T, (0,5,-8)^T]$. A dimensão de S é dois.

No exemplo anterior, as colunas da matriz aumentada que correspondem ao lado esquerdo da equação (1) são os vetores de S. Além disso, o número de linhas (duas) não nulas da forma escalonada é igual á dimensão do $\text{span}(S)$.

Em geral, para problemas em $\mathbb{R}^n$, como no caso do exemplo 43, contamos com a ajuda de um teorema. Observamos que, para uma matriz A, $m \times n$, se considerarmos as linhas como vetores de $\mathbb{R}^n$, então o espaço gerado por esses vetores é chamado o *espaço linha* de A.

Teorema. *Se* $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ *é um conjunto de vetores de* $\mathbb{R}^n$ *e* A *é a matriz construída pondo* $\mathbf{v_1}$ *na linha 1,* $\mathbf{v_2}$ *na linha 2, e assim por diante; e se* B *é a matriz que resulta reduzindo* A *à forma escalonada por linhas, então as linhas não nulas de* B *formam uma base para o espaço linha de* A. *Isto é, as linhas não nulas de* B *formam uma base para* $\text{span}(S)$.

Prova. Seja a matriz $A = \begin{pmatrix} \mathbf{v_1^T} \\ \mathbf{v_2^T} \\ \vdots \\ \mathbf{v_n^T} \end{pmatrix}$.

No processo de escalonar uma matriz por linhas, usamos as operações elementares por linhas. Lembrando do procedimento para realizar es-

sas operações, temos que, se forem obtidas linhas de zeros, essas linhas devem ser combinações lineares das outras linhas do conjunto. As linhas restantes (não nulas) resultam das combinações lineares dos vetores linearmente independentes do conjunto original. Então, o espaço gerado pelas linhas não nulas é o mesmo que o gerado por S. Assim, as linhas não nulas, sendo independentes, formam uma base para $\text{span}(S)$ e $\dim(\text{span}(S))$ número de linhas não nulas.

Exemplo 45. Resolver novamente o Exemplo 43 usando o teorema anterior.

Solução. Construimos A e a escalonamos por linhas,

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 5 & -8 \\ 3 & 2 & -2 \\ 8 & 2 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -\frac{8}{5} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Pelo teorema, $\text{span}(S) = \left[(1,-1,2)^T, \left(0,1,-\frac{8}{5}\right)^T \right]$ e

$\dim(\text{span}(S)) = 2$.

3.6.2 Mudança de Bases

Muitos problemas aplicados podem ser simplificados mudando-se de um sistema de coordenadas para outro. Mudar sistemas de coordenadas em um espaço vetorial é, essencialmente, a mesma coisa que mudar de base. Por exemplo, ao descrever o movimento de uma partícula no plano é muitas vezes conveniente usar uma base de $\mathbb{R}^2$ formada por um vetor tangente unitário T e um vetor normal unitário N, associado à curva, em vez da base canônica $\{\mathbf{e_1}, \mathbf{e_2}\}$.

Nesta seção, vamos discutir o problema de mudar de um sistema de coordenadas para outro.

Vamos mostrar que isso pode ser feito multiplicando-se um vetor de coordenadas dado $\mathbf{x}$ por uma matriz inversível S.

O produto $\mathbf{y} = S\mathbf{x}$ vai ser o vetor de coordenadas para o novo sistema.

3.6.3 Mudanças de Coordenadas em $\mathbb{R}^2$

A base canônica para $\mathbb{R}^2$ é $[\mathbf{e_1}, \mathbf{e_2}]$. Qualquer vetor $\mathbf{x}$ em $\mathbb{R}^2$ pode ser escrito como uma combinação linear dessa base $\mathbf{x} = (x_1, x_2)^T = x_1\mathbf{e_1} + x_2\mathbf{e_2}$.

Os escalares x_1 e x_2 são as coordenadas de $\mathbf{x}$ em relação à base canônica. De fato, para qualquer base $\{\mathbf{u_1}, \mathbf{u_2}\}$ para $\mathbb{R}^2$, um dado vetor $\mathbf{x}$ pode ser representado de maneira única como uma combinação linear $\mathbf{x} = \alpha_1\mathbf{u_1} + \alpha_2\mathbf{u_2}$.

Os escalares α_1, α_2 são as coordenadas de $\mathbf{x}$ em relação à base $\{\mathbf{u_1}, \mathbf{u_2}\}$. Vamos ordenar os elementos da base, de modo que $\mathbf{u_1}$ seja o primeiro vetor da base e $\mathbf{u_2}$ seja o segundo, e vamos denotar a base ordenada por $[\mathbf{u_1}, \mathbf{u_2}]$. Podemos, então, nos referir ao vetor $(\alpha_1, \alpha_2)^T$ como sendo o vetor de coordenadas de $\mathbf{x}$ em relação à base $\{\mathbf{u_1}, \mathbf{u_2}\}$.

Exemplo 46. Sejam $\mathbf{u_1} = (2,1)^T$ e $\mathbf{u_1} = (1,4)^T$. Os vetores são linearmente independentes, e portanto, formam uma base para $\mathbb{R}^2$. O vetor $\mathrm{x} = (7,7)^T$ pode se escrito como uma combinação linear $\mathbf{x} = 3\mathbf{u_1} + \mathbf{u_2}$.

Logo, o vetor de coordenadas de $\mathbf{x}$ em relação a $[\mathbf{u_1}, \mathbf{u_2}]$ é $(3,1)^T$.

Geometricamente, esse vetor no diz como sair da origem e chegar ao ponto $(7,7)$, movendo-nos primeiro na direção de $\mathbf{u_1}$, e depois na direção de $\mathbf{u_2}$.

O vetor de coordenadas de $\mathbf{x}$ em relação à base ordenada $[\mathbf{u_1}, \mathbf{u_2}]$ é $(1,3)^T$. Geometricamente, esse vetor nos diz como sair da origem e chegar a $(7,7)$, movendo-nos primeiro na direção de $\mathbf{u_2}$, e depois na direção de $\mathbf{u_1}$.

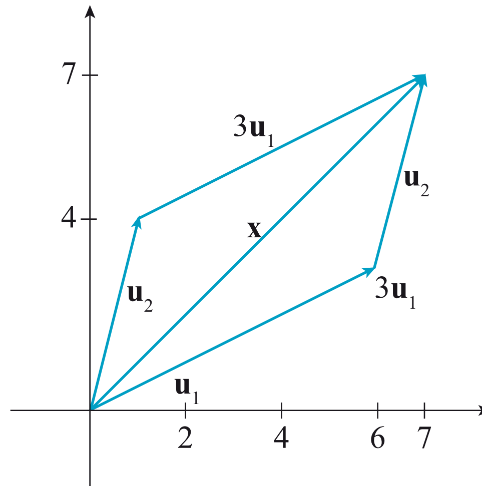


Figura 3.9 - Combinação linear

Uma vez decididos a trabalhar com uma nova base, temos o problema de encontrar as coordenadas em relação a essa nova base. Por exemplo, em vez de usarmos a base canônica para o $\mathbb{R}^2$, usarmos uma base diferente,

$$\mathbf{u_1} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \mathbf{u_2} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

De fato, podemos mudar nos dois sentidos entre os dois sistemas de coordenadas. Vamos considerar os dois problemas seguintes:

1) Dada um vetor $\mathbf{x} = (x_1, x_2)^T$, encontre suas coordenas em relação a $\mathbf{u_1}$ e $\mathbf{u_2}$.

2) Dado um vetor $\mathbf{x} = \alpha_1 \mathbf{u_1} + \alpha_2 \mathbf{u_2}$, encontre suas coordenadas em relação a $\mathbf{e_1}$ e $\mathbf{e_2}$.

Vamos resolver o item 2 primeiro (mais fácil). Para mudar da base $[\mathbf{u_1}, \mathbf{u_2}]$ para a base $[\mathbf{e_1}, \mathbf{e_2}]$, precisamos expressar os elementos da base antiga, $\mathbf{u_1}$ e $\mathbf{u_2}$, em termos dos elementos da nova base, $\mathbf{e_1}$ e $\mathbf{e_2}$.

$$\mathbf{u_1} = 3\mathbf{e_1} + 2\mathbf{e_2}$$
$$\mathbf{u_2} = \mathbf{e_1} + \mathbf{e_2}.$$

Temos, então, que

$$\begin{aligned} \alpha_1 \mathbf{u_1} + \alpha_2 \mathbf{u_2} &= (3\alpha_1 \mathbf{e_1} + 2\alpha_1 \mathbf{e_2}) + (\alpha_2 \mathbf{e_1} + \alpha_2 \mathbf{e_2}) \\ &= (3\alpha_1 + \alpha_2)\mathbf{e_1} + (2\alpha_1 + \alpha_2)\mathbf{e_2}. \end{aligned}$$

Logo, o vetor de coordenadas de $\alpha_1\mathbf{u_1}+\alpha_2\mathbf{u_2}$ em relação a $[\mathbf{e_1},\mathbf{e_2}]$ é $\mathbf{x}=\begin{pmatrix}3\alpha_1+\alpha_2\\2\alpha_1+\alpha_2\end{pmatrix}=\begin{pmatrix}3&1\\2&1\end{pmatrix}\begin{pmatrix}\alpha_1\\\alpha_2\end{pmatrix}$, definindo $U=(\mathbf{u_1},\mathbf{u_2})=\begin{pmatrix}3&1\\2&1\end{pmatrix}$.

Temos que, dado qualquer vetor de coordenadas $\alpha=(\alpha_1,\alpha_2)^T$ em relação a $[\mathbf{u_1},\mathbf{u_2}]$, para encontrar o vetor de coordenadas x correspondentes em relação a $[\mathbf{e_1},\mathbf{e_2}]$, basta multiplicar U e α.

$$\mathbf{x}=U\alpha \tag{1}$$

A matriz U é chamada de matriz de mudança de base de $[\mathbf{u_1},\mathbf{u_2}]$ para $[\mathbf{e_1},\mathbf{e_2}]$.

Para resolver o problema 1, precisamos encontrar a matriz de mudança de base de $[\mathbf{e_1},\mathbf{e_2}]$ para $[\mathbf{u_1},\mathbf{u_2}]$. A matriz U é inversível, já que suas colunas são vetores linearmente independentes.

Da equação (1), temos que: $\alpha=U^{-1}\mathbf{x}$.

Logo, dado um vetor: $\mathbf{x}=(x_1,x_2)^T=x_1\mathbf{e_1}+x_2\mathbf{e_2}$, basta multiplicá-lo por U^{-1} para encontrar seu vetor de coordenadas em relação a $[\mathbf{u_1},\mathbf{u_2}]$. U^{-1} é a matriz de mudança de base de $[\mathbf{e_1},\mathbf{e_2}]$ para $[\mathbf{u_1},\mathbf{u_2}]$.

Exemplo 47. Sejam $\mathbf{u_1}=\begin{pmatrix}3\\2\end{pmatrix},\mathbf{u_2}\begin{pmatrix}1\\1\end{pmatrix}$ e $\mathbf{x}=\begin{pmatrix}7\\4\end{pmatrix}$. Encontre as coordenadas de $\mathbf{x}$ em relação a $[\mathbf{u_1},\mathbf{u_2}]$.

Solução. Pela discussão precedente, a matriz de mudança de base de $[\mathrm{e}_1,\mathrm{e}_2]$ para $[\mathrm{u}_1,\mathrm{u}_2]$ é a inversa de $U=(\mathrm{u}_1,\mathrm{u}_2)=\begin{pmatrix}3&1\\2&1\end{pmatrix}$.

Logo, $\alpha=U^{-1}\mathbf{x}=\begin{pmatrix}1&-1\\-2&3\end{pmatrix}\begin{pmatrix}7\\4\end{pmatrix}=\begin{pmatrix}-3\\2\end{pmatrix}$.

O vetor de coordenadas desejado é $\mathrm{x}=3\mathrm{u}_1-2\mathrm{u}_2$.

Exemplo 48. Seja $\mathbf{b_1}=(1,-1)^T$ e $\mathbf{b_1}=(-2,3)^T$. Encontre a matriz de mudança de base de $[\mathbf{e_1},\mathbf{e_2}]$ para $[\mathbf{b_1},\mathbf{b_2}]$ e determine as coordenadas de $\mathbf{x}=(1,2)^T$ em relação a $[\mathbf{b_1},\mathbf{b_2}]$.

Solução. A matriz de mudança de base de $[\mathbf{b_1},\mathbf{b_2}]$ para $[\mathbf{e_1},\mathbf{e_2}]$ é $B=(\mathbf{b_1},\mathbf{b_2})=\begin{pmatrix}1&-2\\-1&3\end{pmatrix}$.

Logo, a matriz de mudança de base de $[\mathbf{e_1}, \mathbf{e_2}]$ para $[\mathbf{b_1}, \mathbf{b_2}]$ é $B^{-1} = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$.

O vetor de coordenadas de $\mathbf{x}$ em relação a $[\mathbf{b_1}, \mathbf{b_2}]$ é $\alpha = B^{-1}\mathbf{x} = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}$.

E, portanto, $\mathbf{x} = 7b_1 + 3b_2$.

Se $S = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}$ é a matriz de mudança de base de uma base ordenada $[\mathbf{v_1}, \mathbf{v_2}]$ em $\mathbb{R}^2$ para outra base ordenada $[\mathbf{u_1}, \mathbf{u_2}]$, então $\mathbf{v_1} = 1\mathbf{v_1} + 0\mathbf{v_2}$. O vetor de coordenadas de $\mathbf{v_1}$ em relação a $[\mathbf{u_1}, \mathbf{u_2}]$ é dado por $\mathbf{S_1} = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} s_{11} \\ s_{21} \end{pmatrix}$.

Analogamente, para $\mathbf{v_2}, \mathbf{v_2} = 0\mathbf{v_1} + 1\mathbf{v_2}$.

E seu vetor de coordenadas em relação a $[\mathbf{u_1}, \mathbf{u_2}]$ é dado por $\mathbf{S_2} = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} s_{12} \\ s_{22} \end{pmatrix}$.

Logo,

$$\begin{aligned} \mathbf{v_1} &= s_{11}\mathbf{u_1} + s_{21}\mathbf{u_2} \\ \mathbf{v_2} &= s_{12}\mathbf{u_1} + s_{22}\mathbf{u_2}. \end{aligned} \tag{1}$$

Em geral, se os elementos da base antiga $\mathbf{v_1}$ e $\mathbf{v_2}$ são escritos em termos da nova base $[\mathbf{u_1}, \mathbf{u_2}]$, o vetor de coordenadas $\mathbf{s_1} = (s_{11}, s_{21})$ correspondente a $\mathbf{v_1}$ é a primeira coluna da matriz mudança de base S e o vetor de coordenadas $\mathbf{s_2} = (s_{12}, s_{22})$ correspondente a $\mathbf{v_2}$ é a segunda coluna de S. Logo, S é a transposta da matriz de coeficientes em (1).

Exemplo 49. Encontre a matriz de mudança de base de $[\mathbf{v_1}, \mathbf{v_2}]$ para $[\mathbf{u_1}, \mathbf{u_2}]$, onde

$$\mathbf{v_1} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \mathbf{v_2} = \begin{pmatrix} 7 \\ 3 \end{pmatrix}, \mathbf{u_1} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \mathbf{u_2} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Solução. Precisamos escrever $\mathbf{v_1}$ e $\mathbf{v_2}$ em termos dos elementos da nova base $\mathbf{u_1}$ e $\mathbf{u_2}$,

$$\begin{aligned} \mathbf{v_1} &= s_{11}\mathbf{u_1} + s_{21}\mathbf{u_2} \\ \mathbf{v_2} &= s_{12}\mathbf{u_1} + s_{22}\mathbf{u_2}. \end{aligned}$$

A primeira equação pode ser escrita como $\begin{pmatrix} 5 \\ 2 \end{pmatrix} = \begin{pmatrix} 3s_{11} + s_{21} \\ 2s_{11} + s_{21} \end{pmatrix}$.

A solução desse sistema é $(s_{11}, s_{21})^T = (3, -4)^T$. Analogamente, a segunda equação nos leva ao sistema $\begin{pmatrix} 7 \\ 3 \end{pmatrix} = \begin{pmatrix} 3s_{12} + s_{22} \\ 2s_{12} + s_{22} \end{pmatrix}$, cuja solução é $(s_{12}, s_{22})^T = (4, -5)^T$.

Portanto, $S = \begin{pmatrix} 3 & 4 \\ -4 & -5 \end{pmatrix}$ é a matriz de mudança de base de $[\mathbf{v_1}, \mathbf{v_2}]$ para $[\mathbf{u_1}, \mathbf{u_2}]$.

Um método alternativo para mudar de uma base $[\mathbf{v_1}, \mathbf{v_2}]$ para outra base $[\mathbf{u_1}, \mathbf{u_2}]$ é mudar primeiro de $[\mathbf{v_1}, \mathbf{v_2}]$ para a base canônica $[\mathbf{e_1}, \mathbf{e_2}]$ e depois mudar para $[\mathbf{u_1}, \mathbf{u_2}]$.

Dado um vetor $\mathbf{x} \in \mathbb{R}^2$, se $\mathbf{c}$ é o vetor de coordenadas de $\mathbf{x}$ em relação a $[\mathbf{v_1}, \mathbf{v_2}]$ e $\mathbf{d}$ é o vetor de coordenadas de $\mathbf{x}$ em relação a $[\mathbf{u_1}, \mathbf{u_2}]$, então

$$c_1\mathbf{v_1} + c_2\mathbf{v_2} = x_1\mathbf{e_1} + x_2\mathbf{e_2} = d_1\mathbf{u_1} + d_2\mathbf{u_2}.$$

Como V é a matriz de mudança de base de $[\mathbf{v_1}, \mathbf{v_2}]$ para $[\mathbf{e_1}, \mathbf{e_2}]$ e U^{-1} é a matriz mudança de base de $[\mathbf{e_1}, \mathbf{e_2}]$ para $[\mathbf{u_1}, \mathbf{u_2}]$, tem-se que $V\mathbf{c} = \mathbf{x}$ e $U^{-1}\mathbf{x} = \mathbf{d}$.

E, portanto, $U^{-1}V\mathbf{c} = U^{-1}\mathbf{x} = \mathbf{d}$.

Logo , $U^{-1}V$ é a matriz mudança de base de $[\mathbf{v_1}, \mathbf{v_2}]$ para $[\mathbf{u_1}, \mathbf{u_2}]$.

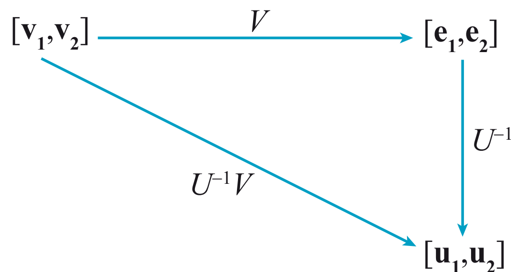


Figura 3.10 - Mudança de base

Exemplo 50. Sejam $[\mathbf{v_1}, \mathbf{v_2}]$ e $[\mathbf{u_1}, \mathbf{u_2}]$ as bases ordenadas do exemplo anterior. A matriz de mudança de base de $[\mathbf{v_1}, \mathbf{v_2}]$ para $[\mathbf{u_1}, \mathbf{u_2}]$ é dada por:

$$U^{-1}V = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix}\begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ -4 & -5 \end{pmatrix}.$$

3.6.4 Mudança de Base em um Espaço Vetorial Geral

Tudo que fizemos até agora pode ser generalizado facilmente para qualquer espaço vetorial de dimensão finita. Vamos começar definindo vetores de coordenadas em um espaço vetorial de dimensão n.

Definição. Seja V um espaço vetorial com base ordenada $E = [\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}]$. Se $\mathbf{v}$ é um elemento qualquer de V, então $\mathbf{v}$ pode ser escrito na forma $\mathbf{v} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n}$, na qual $c_1, c_2, \ldots, c_n$ são escalares. Podemos associar, então, a cada vetor $\mathbf{v}$ um único vetor $\mathbf{c} = (c_1, c_2, ..., c_n)^T$ em $\mathbb{R}^n$. O vetor $\mathbf{c}$, assim definido, é chamado de **vetor de coordenadas** de $\mathbf{v}$ em relação à base ordenada E e é denotado por $[\mathbf{v}]_E$. Os c_i são as **coordenadas** de $\mathbf{v}$ em relação a E.

Os exemplos considerados até agora trataram apenas de mudanças de coordenadas em $\mathbb{R}^2$. Técnicas análogas podem ser usadas em $\mathbb{R}^n$, cujas matrizes de mudança de base serão $n \times n$.

Exemplo 51. Sejam

$$E = [\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}] = [(1,1,1)^T, (2,3,2)^T, (1,3,4)^T]$$
$$F = [\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}] = [(1,1,0)^T, (1,2,0)^T, (1,2,1)^T].$$

Encontre a matriz de mudança de base de E para F. Sejam

$$\mathbf{x} = 3\mathbf{v_1} + 2\mathbf{v_2} - \mathbf{v_3} \text{ e } \mathbf{y} = \mathbf{v_1} - 3\mathbf{v_2} + 2\mathbf{v_3} \text{ ou } [\mathbf{x}]_E = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \text{ e } [\mathbf{y}]_E = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}.$$

Determine as coordenadas de $\mathbf{x}$ e $\mathbf{y}$ em relação à base ordenada F.

Solução. Como no último exemplo, a matriz mudança de base é dada por:

$$U^{-1}V = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 5 \\ 1 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -3 \\ -1 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Os vetores de coordenadas de $\mathbf{x}$ e $\mathbf{y}$ em relação à base ordenada F são dados por:

$$[\mathbf{x}]_E = \begin{pmatrix} 1 & 1 & -3 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 8 \\ -5 \\ 3 \end{pmatrix} \quad \text{e}$$

$$[\mathbf{y}]_F = \begin{pmatrix} 1 & 1 & -3 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix} = \begin{pmatrix} -8 \\ 2 \\ 3 \end{pmatrix}.$$

Você pode verificar que

$$8\mathbf{u}_1 - 5\mathbf{u}_2 + 3\mathbf{u}_3 = 3\mathbf{v}_1 + 2\mathbf{v}_2 - \mathbf{v}_3$$
$$-8\mathbf{u}_1 + 2\mathbf{u}_2 + 3\mathbf{u}_3 = \mathbf{v}_1 - 3\mathbf{v}_2 + 2\mathbf{v}_2.$$

Exemplo 52. Suponha que queremos mudar, em P_2, da base ordenada $[1, x, x^2]$ para a base ordenada $[1, 2x, 4x^2 - 2]$.

Solução. Como $[1, x, x^2]$ é a base canônica para P_2, é mais fácil encontrar a matriz mudança de base de $[1, 2x, 4x^2 - 2]$ para $[1, x, x^2]$.

Como

$$1 = 1\cdot(1) + 0\cdot(x) + 0\cdot(x^2)$$
$$2x = 0\cdot(1) + 2\cdot(x) + 0\cdot(x^2)$$
$$4x^2 - 2 = -2\cdot(1) + 0\cdot(x) + 4\cdot(x^2),$$

a matriz mudança de base é

$$S = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

A inversa de S vai ser a matriz que muda da base $[1, x, x^2]$ para a base $[1, 2x, 4x^2 - 2]$,

$$S^{-1} = \begin{pmatrix} 1 & 0 & 1/2 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{pmatrix}.$$

Dado qualquer $p(x) = ax^2 + bx + c$ em P_2, para encontrar as coordenadas de $p(x)$ em relação a $[1, 2x, 4x^2 - 2]$, basta multiplicar

$$\begin{pmatrix} 1 & 0 & 1/2 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a + c/2 \\ b/2 \\ c/4 \end{pmatrix}.$$

Logo, $p(x) = (c/4)\cdot(x^2) + (b/2)\cdot(x) + (a + c/2)\cdot(1)$. Vimos que a matriz de mudança de base é inversível. De fato, podemos pensar em qualquer matriz inversível como uma matriz de mudança de base.

Em muitos problemas de aplicação, é importante usar o tipo certo de base para o caso em questão. Veremos, em Álgebra Linear II, que a chave para a resolução de quadrados mínimos é usar um tipo especial de base: a base ortonormal. Também vamos considerar um número de aplicações envolvendo autosistemas (autovalores e autovetores) associados a uma matriz $n \times n$. A chave para resolver esse tipo de problema é mudar para uma base para $\mathbb{R}^n$, formada pelos autovetores da matriz.

Exercícios Propostos

1) Indique se os vetores dados no Exercício 1 da página 165 formam ou não base para $\mathbb{R}^2$.

2) Indique se os vetores dados no Exercício 2 da página 165 formam ou não uma base para $\mathbb{R}^3$.

3) Considere os vetores $\mathbf{x_1} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \mathbf{x_2} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \mathbf{x_3} = \begin{pmatrix} 7 \\ -3 \end{pmatrix}$

   a) Mostre que $\mathbf{x_1}$ e $\mathbf{x_2}$ formam uma base para $\mathbb{R}^2$.

   b) Por que $\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}$ têm que ser linearmente dependente?

   c) Qual a dimensão de $[\{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}\}]$?

4) Considere os vetores $\mathbf{x_1} = \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}, \mathbf{x_2} = \begin{pmatrix} -3 \\ 2 \\ 4 \end{pmatrix}, \mathbf{x_3} = \begin{pmatrix} -6 \\ 4 \\ -8 \end{pmatrix}$. Qual a dimensão de $[\{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}\}]$?

5) Considere $\mathbf{x_1} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \mathbf{x_2} = \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix}, \mathbf{x_3} = \begin{pmatrix} 2 \\ 6 \\ 4 \end{pmatrix}$

   a) Mostre que $\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}$ são linearmente dependentes.

   b) Mostre que $\mathbf{x_1}, \mathbf{x_2}$ são linearmente independentes.

   c) Qual a dimensão de $[\{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}\}]$?

   d) Descreva geometricamente $[\{x_1, x_2, x_3\}]$.

6) Alguns dos conjuntos do exercício 2 da página 145 e 146 formam subespaços de $\mathbb{R}^3$. Em cada um desses casos, encontre uma base para o subespaço e determine sua dimensão.

7) Encontre uma base para o subespaço S de $\mathbb{R}^4$ formado por todos os vetores da forma $(a+b, a-b+2c, b, c)^T$, onde a, b, c são números reais. Qual a dimensão de S?

8) Considere os vetores $\mathbf{x_1} = (1,1,1)^T$ e $\mathbf{x_2} = (3,-1,4)^T$.

a) $\mathbf{x_1}$ e $\mathbf{x_2}$ geram $\mathbb{R}^3$? Explique.

b) Seja $\mathbf{x_3}$ um terceiro vetor em $\mathbb{R}^3$. Defina $X = \{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}\}$. Que condição (ou condições) X tem que satisfazer para que $\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}$ formem uma base para $\mathbb{R}^3$?

c) Encontre um terceiro vetor $\mathbf{x_3}$ que estenda o conjunto $\{\mathbf{x_1}, \mathbf{x_2}\}$ a uma base para $\mathbb{R}^3$.

9) Os vetores $\mathbf{x_1} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \mathbf{x_2} = \begin{pmatrix} 2 \\ 5 \\ 4 \end{pmatrix}, \mathbf{x_3} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \mathbf{x_4} = \begin{pmatrix} 2 \\ 7 \\ 4 \end{pmatrix}, \mathbf{x_5} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ geram $\mathbb{R}^3$.

Retire algum (ou alguns) elemento de $\{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}, \mathbf{x_4}, \mathbf{x_5}\}$ de modo a obter uma base para $\mathbb{R}^3$.

10) Seja S o subespaço de P_3 formado por todos os polinômios da forma $ax^2 + bx + 2a + 3b$. Encontre uma base para S.

11) Alguns dos conjuntos do exercício 3 da página 146 formavam subespaços de $\mathbb{R}^{2\times 2}$. Em cada um desses casos, encontre uma base para o subespaço e determine sua dimensão.

12) Para um dos itens a seguir, encontre a matriz que corresponde à mudança de base $[\mathbf{u_1}, \mathbf{u_2}]$ para a base $[\mathbf{e_1}, \mathbf{e_2}]$.

a) $\mathbf{u_1} = (1,1)^T, \mathbf{u_2} = (-1,1)^T$

b) $\mathbf{u_1} = (1,2)^T, \mathbf{u_2} = (2,5)^T$

c) $\mathbf{u_1} = (0,1)^T, \mathbf{u_2} = (1,0)^T$

13) Para cada uma das bases coordenadas $[\mathbf{u_1}, \mathbf{u_2}]$ no exercício 1, encontre a matriz mudança de base de $[\mathbf{e_1}, \mathbf{e_2}]$ para $[\mathbf{u_1}, \mathbf{u_2}]$.

14) Sejam $\mathbf{v_1} = (3,2)^T$ e $\mathbf{v_2} = (4,3)^T$ para cada uma das bases ordenadas $[\mathbf{u_1}, \mathbf{u_2}]$ no exercício 1, encontre a matriz mudança de base de $[\mathbf{v_1}, \mathbf{v_2}]$ para $[\mathbf{u_1}, \mathbf{u_2}]$.

15) Seja $E = [(5,3)^T, (3,2)^T]$ e sejam $\mathbf{x} = (1,1)^T$, $\mathbf{y} = (1,-1)^T$ e $\mathbf{z} = (10,7)^T$. Encontre os vetores de coordenadas $[\mathbf{x}]_E$, $[\mathbf{y}]_E$ e $[\mathbf{z}]_E$.

16) Sejam $\mathbf{u_1} = (1,1,1)^T$, $\mathbf{u_2} = (1,2,2)^T$ e $\mathbf{u_3} = (2,3,4)^T$.

   a) Encontre a matriz mudança de base de $[\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}]$ para $[\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}]$.

   b) Encontre as coordenadas de cada um dos vetores a seguir, em relação a $[\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}]$.

      i) $(3,2,5)^T$ ii) $(1,1,2)^T$ iii) $(2,3,2)^T$

17) Sejam $\mathbf{v_1} = (4,6,7)^T$, $\mathbf{v_2} = (0,1,1)^T$ e $\mathbf{v_3} = (0,1,2)^T$ e sejam $\mathbf{u_1}$, $\mathbf{u_2}$ e $\mathbf{u_3}$ os vetores dados no exercício 16.

   a) Encontre a matriz mudança de base de $[\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}]$ para $[\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}]$.

   b) Se $\mathbf{x} = \mathbf{2v_1} + 3\mathbf{v_2} - 4\mathbf{v_3}$, determine as coordenadas de $\mathbf{x}$ em relação a $[\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}]$.

18) Considere $\mathbf{v_1} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \mathbf{v_2} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, $S = \begin{pmatrix} 3 & 5 \\ 1 & -2 \end{pmatrix}$. Encontre vetores $\mathbf{w_1}$ e $\mathbf{w_2}$ tais que S é a matriz mudança de base de $[\mathbf{w_1}, \mathbf{w_2}]$ para $[\mathbf{v_1}, \mathbf{v_2}]$.

19) Considere $\mathbf{v_1} = \begin{pmatrix} 2 \\ 6 \end{pmatrix}, \mathbf{v_2} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$, $S = \begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix}$. Encontre vetores $\mathbf{u_1}$ e $\mathbf{u_2}$, tais que S é a matriz mudança de base de $[\mathbf{v_1}, \mathbf{v_2}]$ para $[\mathbf{u_1}, \mathbf{u_2}]$.

3.7 Subespaços Associados a Matrizes e Computação de Bases

Formalizaremos alguns dos conceitos colocados em seções anteriores, acrescentando outros subespaços associados com uma matriz.

Definição. Seja A uma matriz $m \times n$.

1) **O espaço nulo** de A, *espnul*(A) é o subespaço de $\mathbb{R}^n$ que consiste nas soluções do sistema linear homogêneo $A\mathrm{x} = 0$.

2) **O espaço linha** de A, $esplin(A)$ é o subespaço de $\mathbb{R}^n$ gerado pelas linhas de A.

3) **O espaço coluna** de A, $espcol(A)$ é o subespaço de $\mathbb{R}^m$ gerado pelas colunas de A.

4) O espaço nulo à esquerda $espnul(A^T)$ é o subespaço de $\mathbb{R}^n$ que consiste nas soluções do sistema linear homogêneo $\mathrm{A}^{\mathrm{T}}x = 0$.

Os quatro espaços definidos acima são classificados os "espaços fundamentais da matriz A".

3.7.1 Uma Base para o Espaço Nulo

Exemplo 53. O espaço nulo da matriz

$$A = \begin{pmatrix} 1 & 3 & 2 & 0 & 1 \\ -1 & -1 & -1 & 1 & 0 \\ 0 & 4 & 2 & 4 & 3 \\ 1 & 3 & 2 & -2 & 0 \end{pmatrix}$$

está formado por todas as soluções do sistema homogêneo $A\mathbf{x} = 0$. Reduzindo esta matriz a sua forma escalonada por linhas, temos:

$$\begin{pmatrix} 1 & 0 & 1/2 & 0 & -1/4 \\ 0 & 1 & 1/2 & 0 & 1/4 \\ 0 & 0 & 0 & 1 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Daqui, fica fácil derivar as soluções do sistema homogêneo:

$$\mathbf{x_1} = -\left(\frac{1}{2}\right)x_3 - \left(\frac{1}{4}\right)x_5$$

$$\mathbf{x_2} = -\left(\frac{1}{2}\right)x_3 - \left(\frac{1}{4}\right)x_5$$

$\mathbf{x_3}$ é qualquer um

$$\mathbf{x_4} = -\left(\frac{1}{2}\right)x_5$$

$\mathbf{x_5}$ é qualquer um.

Logo, $espnul(A)$ consiste de vetores da forma:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -\frac{x_3}{2} - \frac{x_5}{4} \\ -\frac{x_3}{2} - \frac{x_5}{4} \\ -\frac{x_5}{2} \\ x_5 \end{pmatrix} = x_3 \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -\frac{1}{4} \\ -\frac{1}{4} \\ 0 \\ -\frac{1}{2} \\ 1 \end{pmatrix} = x_3\mathbf{u} + x_5\mathbf{v} .$$

Já que x_3 e x_5 são arbitrários, temos que $espnul(A) = \text{span}\{\mathbf{u}, \mathbf{v}\}$. Além disso, os seguintes argumentos mostram que $\mathbf{u}$ e $\mathbf{v}$ são linearmente independentes. *Para ter* $x_3\mathbf{u} + x_5\mathbf{v} = 0$, a terceira e a quinta coordenadas devem ser nulas, isto é, $x_3 = x_5 = 0$. Em conseqüência, $\mathbf{u}$ e $\mathbf{v}$ formam uma base de $espnul(A)$.

O que temos feito neste exemplo se aplica ao caso mais geral. Podemos esboçar um procedimento para achar a base do espaço nulo.

Definição. Chamamos a nulidade de A a dimensão do espaço nulo de A, isto é, nulidade de $A = \dim(espnul(A))$.

Para o exemplo anterior, a $\dim(espnul(A)) = 2$.

Se a solução geral de $A\text{x} = 0$ é $\mathbf{x} = c_1\mathbf{v_2} + c_2\mathbf{v_2} + \ldots + c_k\mathbf{v_k}$, onde $c_1, c_2, \ldots, c_k$ são as variáveis livres, temos que $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ formam uma base para $espnul(A)$.

As bases dos núcleos das transformações lineares podem ser achadas de uma maneira similar a nulidade. Isto é um assunto a ser tratado no próximo capítulo.

3.7.2 Uma Base para o Espaço Linha

A partir do seguinte teorema, teremos um procedimento padrão para determinar a base do espaço linha de uma matriz.

Teorema. *Duas matrizes equivalentes por linhas têm o mesmo espaço linha.*

Logo, a forma escalonada por linhas de uma matriz fornece-nos o espaço linha.

Do exemplo anterior, temos que a forma escalonada por linhas:

$$\begin{pmatrix} 1 & 0 & 1/2 & 0 & -1/4 \\ 0 & 1 & 1/2 & 0 & 1/4 \\ 0 & 0 & 0 & 1 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Logo, como os vetores linha (não nulos) são linearmente independentes, formam uma base e

$$esplin(A) = span\left\{\begin{pmatrix} 1 \\ 0 \\ \frac{1}{2} \\ 0 \\ -\frac{1}{4} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \frac{1}{2} \\ 0 \\ \frac{1}{4} \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \frac{1}{2} \end{pmatrix}\right\}.$$

Definição. O posto de uma matriz A, $\text{posto}(A)$, é a dimensão de seu espaço linha.

No exemplo anterior, temos que $\dim(esplin(A)) = \text{posto}(A) = 3$.

3.7.3 Uma Base para o Espaço Coluna

Exemplo 54. O espaço coluna da matriz

$$A = \begin{pmatrix} 1 & 3 & 2 & 0 & 1 \\ -1 & -1 & -1 & 1 & 0 \\ 0 & 4 & 2 & 4 & 3 \\ 1 & 3 & 2 & -2 & 0 \end{pmatrix}$$

é gerado pelos cinco vetores coluna. Levando à forma escalonada por linhas, temos:

$$A = \begin{pmatrix} 1 & 3 & 2 & 0 & 1 \\ 0 & 2 & 1 & -1 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Já que [col 3] e [col 5] não têm pivô, a [col 5] é uma combinação linear de [col 1] e [col 4], e a [col 3] é a uma combinação linear de [col 1] e [col 2], temos que o *espaço coluna de* A:

Col[A] = span{[col 1], [col 2], [col 3], [col 4], [col 5]}
= span{[col 1], [col 2], [col 3], [col 4]}
= span{[col 1], [col 2], [col 4]}.

Por outro lado, se mantemos apenas [col 1], [col 2], [col 4], então

$$\begin{pmatrix} 1 & 3 & 0 \\ -1 & -1 & 1 \\ 0 & 4 & 4 \\ 1 & 3 & -2 \end{pmatrix},$$

cuja forma escalonada é

$$\begin{pmatrix} 1 & 3 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix},$$

onde todas as colunas possuem pivô. Então, já que essas três colunas são linearmente independentes e geram col(A), concluímos que

$$espcol(A) = span\left\{\begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 4 \\ 3 \end{pmatrix}\right\}, \begin{pmatrix} 0 \\ 1 \\ 4 \\ -2 \end{pmatrix}.$$

O que temos neste exemplo pode ser aplicado a qualquer matriz.

1) As colunas sem pivô são combinação linear das colunas que a precedem. Assim, elas podem ser eliminadas sem afetar o espaço coluna.
2) As colunas com pivô são linearmente independentes.

E nos conduz a um procedimento padrão para determinar uma base para o espaço coluna. Para uma matriz A, as colunas pivôs de A formam uma base de $espcol(A)$.

Frisamos que são as colunas de A, e não as colunas da forma escalonada por linhas, que formam a base.

Veja como este teorema se aplica no exemplo anterior, comparando as dimensões dos espaços envolvidos.

Teorema. *Seja A uma matriz $m \times n$, então a dimensão do espaço linha é igual à dimensão do espaço coluna.*

3.8 Espaços Linha/Coluna e os Sistemas Lineares

Os conceitos de espaço linha e espaço coluna são úteis no estudo de sistemas lineares. O sistema $A\mathbf{x}=\mathbf{b}$ pode se escrito na forma

$$A\mathbf{x}=\begin{pmatrix} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n \\ \vdots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n \end{pmatrix}=x_1\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}+x_2\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}+\ldots+x_n\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$$

$$=x_1\mathrm{a}_1+x_2\mathrm{a}_2+\ldots+x_n\mathrm{a}_n=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Segue-se que o sistema $A\mathbf{x}=\mathbf{b}$ é compatível se, e somente se, $\mathbf{b}$ pode ser escrito como uma combinação linear dos vetores colunas de A. Temos, então, as seguintes caracterizaçações de sistemas compatíveis:

- $A\mathbf{x}=\mathbf{b}$ é compatível se, e somente se, $\mathbf{b}$ pertence ao espaço coluna de A.

Fazendo $\mathbf{b}=0$, temos que $A\mathbf{x}=0$, então temos:

- **O sistema $A\mathbf{x}=0$ tem apenas a solução trivial $\mathbf{x}=0$ se, e somente se, os vetores colunas de A são linearmente independentes.**

Podemos inferir que, para uma matriz $n\times n$:

- **Uma matriz A é inversível se, e somente se, os vetores colunas de A formam uma base para $\mathbb{R}^n$.**

Se A é uma matriz $m\times n$, então a soma do posto de A com a nulidade de A é igual a n.

$$\text{posto}(A)+\dim(espnul(A))=n\,.$$

Exemplo 55. O espaço coluna de

$$A = \begin{pmatrix} 1 & 3 & 2 & 0 & 1 \\ -1 & -1 & -1 & 1 & 0 \\ 0 & 4 & 2 & 4 & 3 \\ 1 & 3 & 2 & -2 & 0 \end{pmatrix}$$

é gerado pelos cinco vetores coluna. Os cinco vetores geradores podem ser simplificados utilizando as operações elementares por coluna para produzir a forma escalonada por colunas:

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \end{pmatrix}.$$

Já que as colunas nulas não contribuem para gerar o espaço, temos: $espcol(A) = espcol(B) = span$ (primeiras três colunas de B).

Além disso, as três primeiras colunas de B são linearmente independentes. Logo, as colunas não nulas de B *são:*

$$[\text{col } 1] \text{ de } B = (1, -1, 0, 1)$$

$$[\text{col } 2] \text{ de } B = (0, 1, 2, 0)$$

$$[\text{col } 3] \text{ de } B = (0, 0, 1, -1)$$

e formam uma base de $espcol(A)$. Agora $espcol(B) = espcol(A)$.

O que temos feito pode ser aplicado a qualquer matriz A. Em resumo, temos:

1) Se B é a forma escalonada de A, então $espcol(B) = espcol(A)$.
2) As colunas não nulas de B são linearmente independentes e geram $\text{col}(B)$.

Atenção: frisamos que são as colunas da forma escalonada por colunas as que determinam esta base, e não as colunas de A.

Para uma matriz A, as colunas não nulas da forma escalonada por colunas formam uma base para $espcol(B)$.

Sugestão. Refaça o exemplo 55 utilizando esta técnica.

Exercícios Propostos

1) Para cada uma das matrizes a seguir, encontre uma base para o espaço linha, uma base para o espaço coluna e uma base para o núcleo.

a) $\begin{pmatrix} 1 & 3 & 2 \\ 2 & 1 & 4 \\ 4 & 7 & 8 \end{pmatrix}$

b) $\begin{pmatrix} -3 & 1 & 3 & 4 \\ 1 & 2 & -1 & -2 \\ -3 & 8 & 4 & 2 \end{pmatrix}$

c) $\begin{pmatrix} 1 & 3 & -2 & 1 \\ 2 & 1 & 3 & 2 \\ 3 & 4 & 5 & 6 \end{pmatrix}$

2) Em cada um dos itens a seguir, determine a dimensão do subespaço de $\mathbb{R}^3$ gerado pelos vetores dados.

a) $\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 4 \end{pmatrix}, \begin{pmatrix} -3 \\ 3 \\ 6 \end{pmatrix}$

b) $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$

c) $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$

3) Seja $A = \begin{pmatrix} 1 & 2 & 2 & 3 & 1 & 4 \\ 2 & 4 & 5 & 5 & 4 & 9 \\ 3 & 6 & 7 & 8 & 5 & 9 \end{pmatrix}$.

a) Calcule a forma escalonada reduzida por linhas U de A. Quais os vetores colunas de U que correspondem às variáveis livres? Escreva cada um desses vetores colunas como uma combinação linear dos vetores colunas correspondentes às variáveis líderes.

b) Quais os vetores colunas de A que correspondem às variáveis líderes de U? Esses vetores colunas formam uma base para o espaço coluna de A. Escreva cada um dos vetores colunas de A como uma combinação linear dos vetores dessa base.

4) Para cada uma das escolhas de A e b a seguir, determine se b pertence ao espaço coluna de A e diga se o sistema $A\mathrm{x} = b$ é ou não compatível.

a) $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, $b = \begin{pmatrix} 4 \\ 8 \end{pmatrix}$

b) $A = \begin{pmatrix} 3 & 6 \\ 1 & 2 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

c) $A = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$, $b = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$

d) $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

e) $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 2 \\ 5 \\ 2 \end{pmatrix}$

f) $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \end{pmatrix}$, $b = \begin{pmatrix} 5 \\ 10 \\ 5 \end{pmatrix}$

5) Para cada um dos sistemas compatíveis no Exercício 4, examine os vetores colunas da matriz de coeficientes para determinar se o sistema tem uma solução ou uma infinidade de soluções.

6) Quantas soluções o sistema $A\mathbf{x} = \mathbf{b}$ vai ter se $\mathbf{b}$ pertencer ao espaço coluna de A e se os vetores colunas de A forem linearmente independentes? Explique.

7) Seja A uma matriz $m \times n$ com $m > n$. Seja $b \in \mathbb{R}^m$ e suponha que $espnul(A) = \{0\}$.

   a) O que você pode concluir sobre os vetores colunas de A? Eles são linearmente independentes? Eles geram $\mathbb{R}^m$? Explique.

   b) Quantas soluções o sistema $A\mathbf{x} = \mathbf{b}$ vai ter se $\mathbf{b}$ não pertencer ao espaço coluna de A? Quantas soluções o sistema vai ter se b pertencer ao espaço coluna de A? Explique.

8) Sejam A e B matrizes 6×5. Se $\dim(espnul(A)) = 2$, qual o posto de A? Se o posto de B for 4, qual vai ser a $\dim(espnul(B))$?

Resumo

O conceito e as principais propriedades dos espaços vetoriais foram definidos e desenvolvidos ao longo deste capítulo.

Alguns, tais como E^n (espaço euclidiano, n-dimensional), foram obtidos como generalizações diretas de espaços de duas e três dimensões; entretanto, outros, tais como espaços dos polinômios, das funções, ou das matrizes, surgiram por problemas que aparecem

em cálculo, equações diferenciais e na matemática aplicada. Não obstante que, nos exemplos, cada um dos conjuntos considerados tenha diferente natureza, os espaços de dimensão finita têm características comuns, tais como as propriedades, a estrutura dos subespaços, bases, dimensão, etc.

Tópicos que serão estudados futuramente.

O problema da determinação da base de um espaço vetorial poder ser visto como um dos problemas fundamentais da álgebra linear. A compreensão deste conceito se tornará extremamente importante quando for discutido o problema de diagonalização das matrizes. Sua importância prática encontra-se na análise de problemas aplicados.

O teorema fundamental deste capítulo garantiu a existência de uma base para todo espaço de dimensão finita. De fato, a prova demonstrou que uma base pode ser construída por qualquer conjunto de vetores que gerassem V.

Para uma base dada, a representação de um vetor x em V é única; os coeficientes da combinação linear de vetores da base são chamados as coordenadas de $\mathbf{x}$ nessa base.

As bases ortonormais têm boas propriedades, que serão exploradas na disciplina Álgebra Linear II.

As bases para um espaço vetorial não são únicas. Se forem dadas duas bases para um mesmo espaço vetorial V e para um vetor x que está em V, as coordenadas de x, em relação a ambas as bases, estão relacionadas pela matriz de mudança de base. As matrizes de mudança de base serão usadas nas representações de funções importantes, chamadas transformações lineares.

Em Cálculo, foi necessário introduzir a estrutura dos números reais antes de desenvolver o conceito de uma função (domínios e imagens eram subconjuntos de $\mathbb{R}$).

Já que as aplicações lineares que estudaremos no próximo capítulo possuem seus respectivos domínios e imagens como sendo espaços vetoriais, os assuntos estudados neste capítulo são indispensáveis para introduzir o conceito de transformação linear.

Bibliografia Comentada

LAY, David C. *Álgebra linear e suas aplicações*. 2. ed. [S.l]: LTC, [200-?].

O texto fornece uma introdução elementar e moderna da álgebra linear e algumas de suas aplicações interessantes, acessível a alunos com a maturidade que dois semestres completos de matemática em nível de terceiro grau, em disciplinas de cálculo em geral, lhes conferem. O objetivo é ajudar os alunos a dominar os conceitos e habilidades básicos que usarão mais tarde em suas carreiras. Os tópicos escolhidos seguem as recomendações do Linear Algebra Curriculum Study Group, que, por sua vez, baseiam-se em uma cuidadosa pesquisa sobre as necessidades reais dos alunos e em um consenso entre os profissionais dos muitos campos que usam a Álgebra Linear.

POOLE, David. *Álgebra linear*. São Paulo: Pioneira Thomson Learning, 2004.

Este livro foi estruturado de forma bastante flexível, com a preocupação central de que a álgebra linear constitua um assunto estimulante o suficiente e de fácil aprendizado, tornando-a mais acessível ao estudante. Escrito de forma clara, direta e objetiva, aborda temas como Vetores, Matrizes, Autovalores e Autovetores, Ortogonalidade, Espaços Vetoriais, e Distância e Aproximação. A apresentação de conceitos-chave com antecedência, a ênfase em vetores e geometria e os inúmeros exercícios e exemplos que reforçam o fato de a Álgebra Linear ser uma ferramenta valiosa para a modelagem de problemas da vida real consistem no principal diferencial deste livro. A apresentação de pequenos esboços biográficos de muitos dos matemáticos que contribuíram para o desenvolvimento da Álgebra Linear é outro diferencial, valorizando a história da matemática.

STEVE, Leon J. *Álgebra linear com aplicações*. 4. ed. [S.l.]: LTC, [200-?].

Este livro é apropriado para alunos que tenham conceitos básicos de matrizes e tenham passado por um curso de Geometria Analítica. O estudante deve estar também familiarizado com as noções básicas de Cálculo Diferencial e Integral. Esta nova edição, ao mesmo tempo que mantém a essência das edições anteriores, incorpora uma série de melhorias substanciais: - Conjunto de Exercícios Computacionais em cada Capítulo; - Mais Motivação Geométrica; - Nova Aplicação Envolvendo Teoria dos Grafos e Redes; - Motivação Adicional para a Definição de Determinantes; - A seção sobre Mudança de Base foi transferida para o Cap. 3; - Revisões Importantes na seção sobre Espaços Munidos de Produto Interno; - A seção sobre Normas Matriciais foi transferida para o Cap. 7; - Nova Aplicação: Aproximação de Funções por Polinômios Trigonométricos; - Revisões no Cap. 6.

Capítulo 4

Transformações Lineares

Capítulo 4

Transformações Linares

Este capítulo tem como objetivos principais introduzir a noção de transformação linear e mostrar as relações que existem entre as transformações lineares e as matrizes.

Começamos introduzindo a definição de transformação linear e apresentando exemplos que ilustram os efeitos geométricos de algumas transformações no plano. A seguir, mostramos a associação que existe entre matrizes e transformações lineares, isto é, que toda transformação linear T de um espaço vetorial V de dimensão n num espaço vetorial W de dimensão m pode ser representada por uma matriz $A_{m\times n}$. Esta última ideia nos permite considerar a relação entre as diferentes matrizes que representam o mesmo operador linear. Em muitas aplicações é desejável uma base específica de modo que a matriz que representa a transformação linear seja diagonal ou tenha alguma outra forma simples.

4.1 Introdução

Uma transformação linear é uma aplicação que leva vetores de um espaço vetorial em outro espaço vetorial.

Denotaremos uma transformação linear como $T: V \to W$, onde T é a *transformação linear* (uma aplicação, mapeamento, função, etc.) de V em W, onde V (um espaço vetorial) é o *domínio* e W (um espaço vetorial) é o *contradomínio*. Para $\mathbf{x} \in \mathbb{R}^m$, o vetor $T(\mathbf{x}) \in \mathbb{R}^m$ é chamado de *imagem* de $\mathbf{x}$ (sob a ação de T).

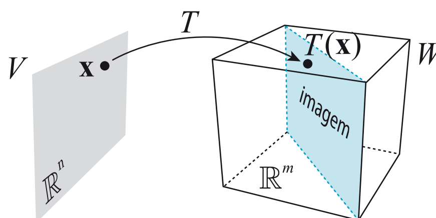


Figura 4.1 – Domínio, contradomínio e imagem de $T: \mathbb{R}^n \to \mathbb{R}^m$

Já encontramos uma ferramenta matemática que nos permitiu converter vetores. Se considerarmos a matriz $A_{m \times n}$ e $\mathbf{x} \in \mathbb{R}^n$, então fazendo o produto $A\mathbf{x}$ obtemos um novo vetor $\mathbf{y} \in \mathbb{R}^m$, definido como $\mathbf{y} = A\mathbf{x}$.

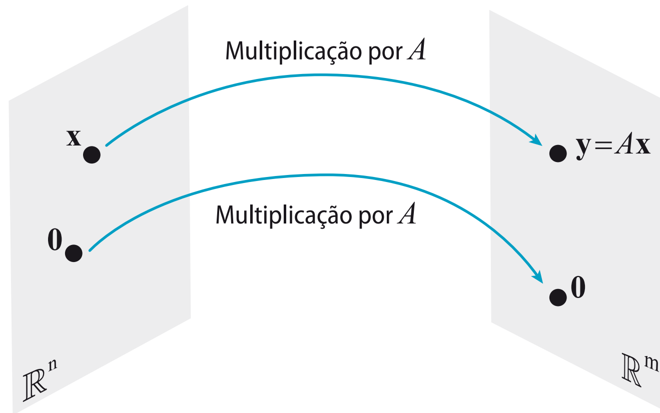


Figura 4.2 – Transformando vetores via multiplicação de matrizes

Definição. Sejam V e W dois espaços vetoriais. Uma Transformação Linear (TL) é uma função de V em W que satisfaz as seguintes condições:

1) Quaisquer que sejam $\mathbf{u}$ e $\mathbf{v}$ em V, $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$;
2) Quaisquer que sejam $\lambda \in \mathbb{R}$ e $\mathbf{u}$ em V, $T(\lambda\mathbf{v}) = \lambda T(\mathbf{v})$;
3) Ou equivalentemente (1) e (2) podem ser resumidas com $\lambda \in \mathbb{R}, T(\lambda\mathbf{u}+\mathbf{v}) = \lambda T(\mathbf{u}) + T(\mathbf{v})$.

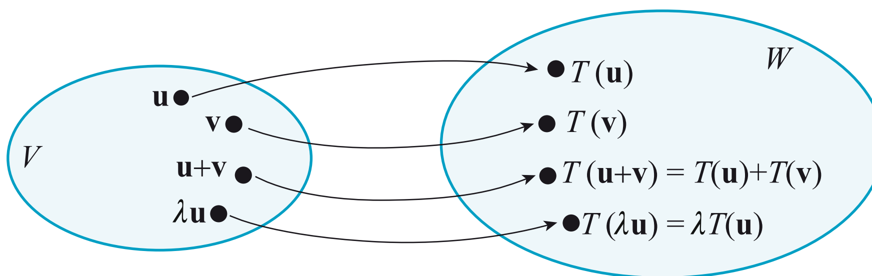


Figura 4.3 – Definição de uma Transformação Linear

Observamos da definição que uma TL "preserva" as operações de adição entre vetores e a multiplicação por escalar.

Exemplo 1. Seja $V = W = \mathbb{R}$, isto é, $T: \mathbb{R} \to \mathbb{R}$, onde a transformação está definida como $T(x) = \alpha x$.

Essa transformação pode ser considerada das mais elementares: T é simplesmente a função linear $f(x) = \alpha x$ e, neste caso, os vetores de V e W são escalares. Temos que:

$$T(\boldsymbol{u}+\boldsymbol{v}) = T(x+y) = \alpha(x+y) = \alpha x + \alpha y = T(x) + T(y) = T(\boldsymbol{u}) + T(\boldsymbol{v}) \text{ e}$$

$$T(\lambda \boldsymbol{u}) = T(\lambda x) = \lambda \alpha x = \lambda(\alpha x) = \lambda T(x) = \lambda T(\boldsymbol{u}),$$

que verificam as condições de TL.

O nome transformação linear certamente foi inspirado nesse caso em que $V = W = \mathbb{R}$; o gráfico de $f(x) = \alpha x$ é uma reta passando pela origem.

Exemplo 2. Consideremos $V = \mathbb{R}^2$ e $w = \mathbb{R}^3$ e $T: \mathbb{R}^2 \to \mathbb{R}^3$ definida como: $(x, y) \to (2x, 0, x+y)$ ou $T(x, y) \to (2x, 0, x+y)$

Dados $\mathbf{u}$, $\mathbf{v} \in \mathbb{R}^2$, com $\mathbf{u} = (x_1, y_1)$ e $\mathbf{v} = (x_2, y_2)$.

Usaremos a definição para determinar se T é uma transformação linear.

1) $T = (\mathbf{u} + \mathbf{v}) = T(x_1, y_1) + (x_2, y_2) =$

$= T(x_1, x_2, y_1, y_2) =$

$= (2(x_1, x_2), 0, (x_1, x_2) + (y_1, y_2)) =$

$= (2x_1, 0, x_1, x_2) + 2(x_2, 0, x_2 + y_2) =$

$T = (x_1, y_1) + T(x_2, y_2) =$

$T = (\mathbf{u}) + T(\mathbf{v})$

Assim, a primeira condição é satisfeita. Continuando, iremos conferir a segunda condição.

$T(\lambda \mathbf{u}) =$

$= T(\lambda(x, y) =$

$= T(\lambda x, \lambda y) =$

$= (2\lambda x, 0, x + y) =$

$= \lambda T(x + y) =$

$= \lambda T(\mathbf{u})$

Esta também é satisfeita, então T é uma transformação linear.

Contraexemplo. A aplicação $T : \mathbb{R} \to \mathbb{R}$, $T(x) = x^2$ *não é* uma transformação linear, pois $T(x + y) = (x + y)^2 = x^2 + 2xy + y^2$ e por outro lado $T(x) + T(y) = x^2 + y^2$.

Portanto, T é uma transformação linear desde que

$T(x + y) \neq T(x) + T(y)$.

4.1.1 Transformações Lineares do Plano no Plano

Agora iremos apresentar uma visão geométrica das TL, dando alguns exemplos de transformações no plano, isto é, transformações do tipo $T : \mathbb{R}^2 \to \mathbb{R}^2$. Você verá assim que uma expansão, por exemplo, uma rotação e certas deformações podem ser descritas por transformações lineares.

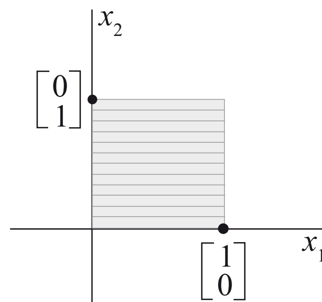


Figura 4.4 - O quadrado unitário

a) Expansão (ou Contração Uniforme)

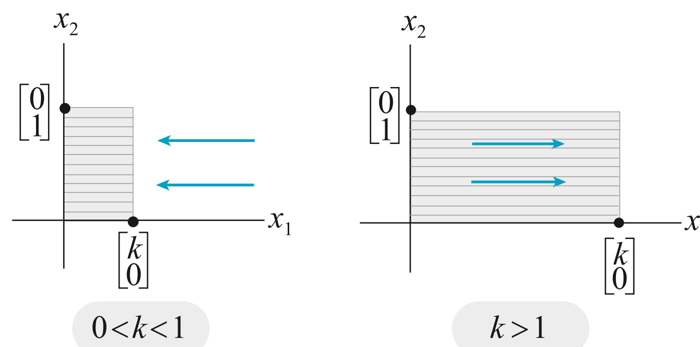
Uma expansão (dilatação) ou contração (compressão) é uma TL do tipo $T = (\mathbf{u}) = k\mathbf{u}$, $k \in \mathbb{R}$.

Exemplo 3. Considere a transformação $T(\mathrm{x}) = 2\mathrm{x} = 2\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$, para todo $\mathrm{x} \in \mathbb{R}^2$.

Como $T(\mathrm{x}+\mathrm{y}) = 2(\mathrm{x}+\mathrm{y}) = 2\mathrm{x}+2\mathrm{y} = T(\mathrm{x})+T(\mathrm{y})$ e $T(\lambda \mathrm{x}) = 2(\lambda \mathrm{x}) = \lambda(2\mathrm{x}) = \lambda T(\mathrm{x})$, assim temos uma TL. Podemos pensar no efeito que T produz, como "esticando" cada vetor por um fator de 2, isto é, modificando apenas o módulo.

Expansão ou contração horizontal

$(x_1, x_2)^T \to (kx_1, x_2)^T$



Expansão ou contração vertical

$(x_1, x_2)^T \to (x_1, kx_2)^T$

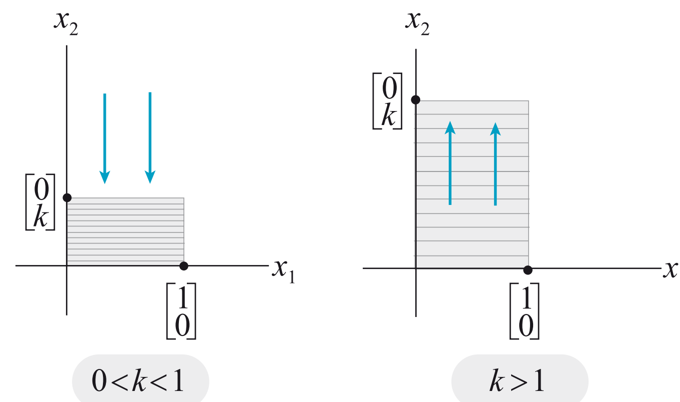


Figura 4.5 – Expansões e contrações

b) Projeção (sobre o eixo $\overrightarrow{OX}$)

Exemplo 4. Seja a TL definida por $T = (\mathbf{x}) = x_1 e_1$.

Se $\mathrm{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, temos $T = (\mathbf{x}) = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$.

Temos uma transformação linear que projeta todos os vetores do plano sobre o eixo das abscissas.

- Projeção no eixo x_1 do vetor $(x_1, x_2)^T$ $(x_1, x_2)^T \rightarrow (x_1, 0)^T$
- Projeção no eixo x_2 do vetor $(x_1, x_2)^T$ $(x_1, x_2)^T \rightarrow (0, x_2)^T$

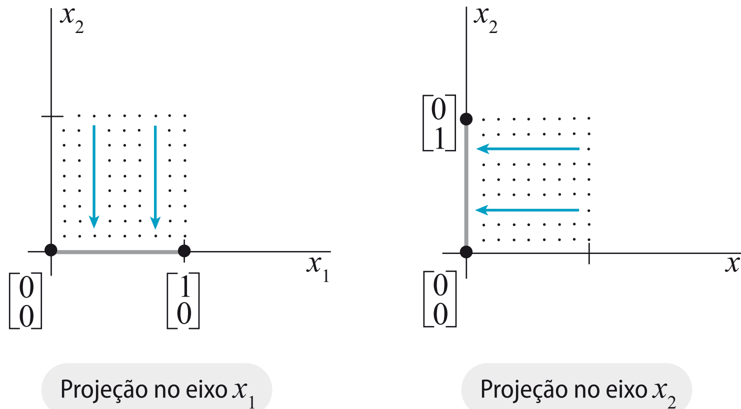


Figura 4.6 – Projeções do quadrado unitário

c) Reflexões

Em relação ao eixo $\overrightarrow{OX}$

Exemplo 5. Considere a aplicação definida por $T = (\mathbf{x}) = (x_1, -x_2)^T$ para todo $\mathbf{x} \in \mathbb{R}^2$.

Este operador reflete um vetor $\mathbf{x}$ em relação ao eixo das abscissas.

$$(x_1, x_2)^T \rightarrow (x_1, -x_2)^T$$

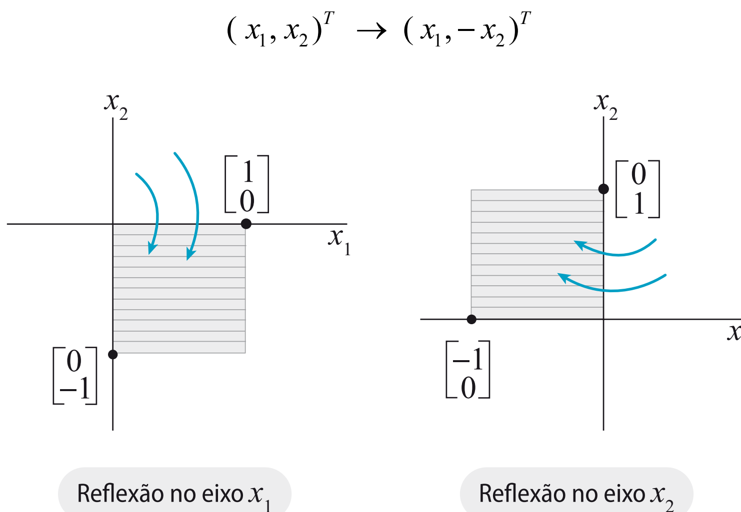


Figura 4.7 – Reflexões

Em relação à origem

Exemplo 6. A aplicação definida por $T(\mathbf{x}) = (x_1, -x_2)^T$ para todo $\mathbf{x} \in \mathbb{R}^2$ inverte os vetores em torno da origem.

$$(x_1, x_2)^T \to (-x_1, -x_2)^T$$

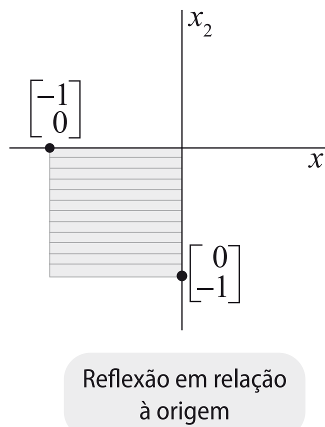


Figura 4.8 – Reflexão em relação à origem

Em relação à reta $y = -x$ $(x_2 = -x_1)$

A aplicação definida por $T(\mathbf{x}) = (-x_2, -x_1)^T$ para todo $\mathbf{x} \in \mathbb{R}^2$ inverte os vetores em torno da reta $y = -x$.

$$(x_1, x_2)^T \to (-x_2, -x_1)^T$$

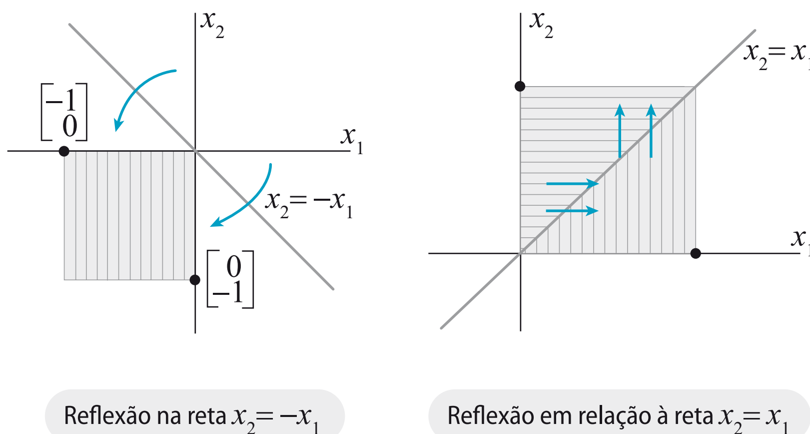


Figura 4.9 – Reflexão em relação às retas $x_2 = -x_1$ e $x_2 = x_1$

d) Rotações

- O operador $T(\mathbf{x}) = (-x_2, x_1)^T$ roda cada vetor $\mathbf{x} \in \mathbb{R}^2$ em 90º em torno da origem, no sentido anti-horário.

- Podemos encontrar uma aplicação que generaliza o caso anterior: a rotação de um vetor em um ângulo θ qualquer (com o sentido de rotação predeterminado). A transformação pode ser conseguida considerando as seguintes relações:

$$y_1 = r\cos(\alpha + \theta) = r(\cos\alpha\cos\theta - \operatorname{sen}\alpha\operatorname{sen}\theta).$$

Porém, $r\cos\alpha = x_1$ e $r\operatorname{sen}\alpha = x_2$.

Então, $y_1 = x_1\cos\theta - x_2\operatorname{sen}\theta$.

Analogamente,

$$y_2 = r\operatorname{sen}(\alpha + \theta) = r(\operatorname{sen}\alpha\operatorname{sen}\theta + \cos\alpha\cos\theta) = x_2\cos\theta + x_1\operatorname{sen}\theta.$$

Logo, a aplicação $T(\mathbf{x}) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1\cos\theta - x_2\operatorname{sen}\theta \\ x_1\operatorname{sen}\theta + x_2\cos\theta \end{pmatrix}$ descreve a rotação de um vetor em um ângulo θ (nesse caso, no sentido anti-horário).

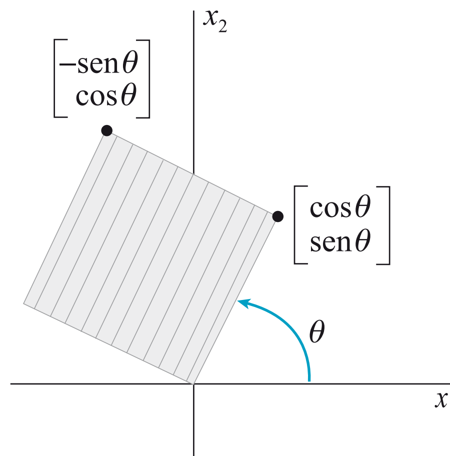


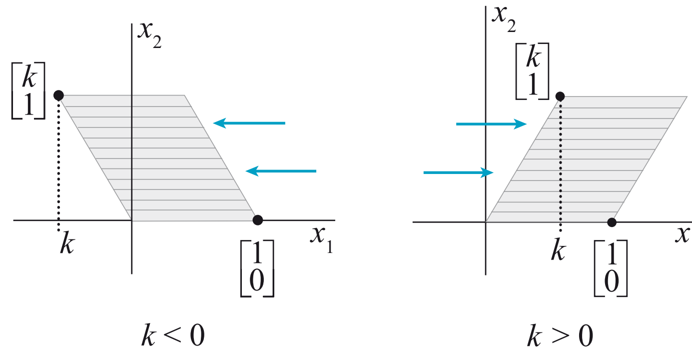
Figura 4.10 – Rotação anti-horária em um ângulo θ

Se considerarmos o caso particular com $\theta = \dfrac{\pi}{2}$, obteremos como resultado novamente $T(\mathbf{x}) = (-x_2, x_1)^T$.

e) Cisalhamento Horizontal

O cisalhamento horizontal é dado pela relação $T(\mathbf{x}) = \begin{pmatrix} x_1 + ax_2 \\ x_2 \end{pmatrix}$. Ele consiste na modificação da primeira coordenada do vetor $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

Cisalhamento horizontal



Cisalhamento vertical

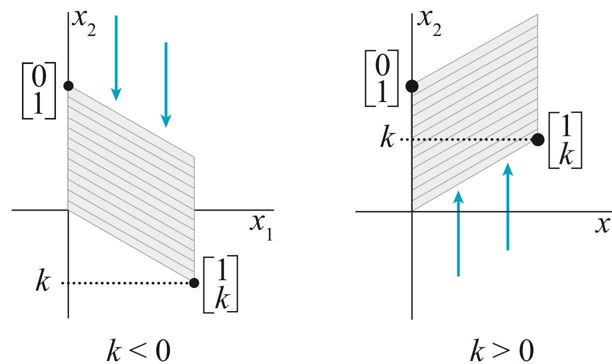


Figura: 4.11 - Cisalhamentos

f) Translação

Considere o vetor $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ e a seguinte aplicação:

$$T(\mathbf{x}) = \mathbf{x} + a = \begin{pmatrix} x_1 + a_1 \\ x_2 + a_2 \end{pmatrix}.$$

Esta *não é* uma transformação linear a não ser que a seja o vetor nulo.

De fato:

$$T(\mathbf{x}+\mathbf{y}) = \begin{pmatrix} (x_1+y_1)+a_1 \\ (x_2+y_2)+a_2 \end{pmatrix}$$
$$= \begin{pmatrix} x_1+a_1 \\ x_2+a_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$
$$= T(\mathbf{x})+\mathbf{y} \neq T(\mathbf{x})+T(\mathbf{y}) .$$

Logo veremos que o uso de uma propriedade das TL simplificará a demonstração desse último exemplo. (Antecipando: a imagem do vetor nulo deve ser o vetor nulo).

4.1.2 Transformações Lineares de $\mathbb{R}^n$ em $\mathbb{R}^m$

Exemplo 7. A aplicação $T:\mathbb{R}^2 \to \mathbb{R}^3$ definida por $T(\mathbf{x}) = (x_2, x_1, x_1+x_2)^T$ é linear, pois

$$T(\mathbf{x}+\mathbf{y}) = (x_2+y_2, x_1+y_1, x_1, y_1, x_2, y_2)^T$$
$$= (x_2, x_1, x_1+x_2)^T + (y_2, y_1, y_1+y_1)^T$$
$$= T(\mathbf{x})+T(\mathbf{y})$$

e

$$T(\lambda \boldsymbol{x}) = (\lambda x_2, \lambda x_1, \lambda x_1, +\lambda x_2)^T$$
$$= \lambda(x_2, x_1, x_1, +x_2)^T$$
$$= \lambda T(\boldsymbol{x}) .$$

Este exemplo nos permite introduzir uma relação entre matrizes e transformações lineares.

Se definirmos a matriz A como $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}$, então, com $\mathbf{x} \in \mathbb{R}^2$ temos:

$$A\mathbf{x} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \\ x_1+x_2 \end{pmatrix} = (x_2, x_1, x_1+x_2)^T = T(\mathbf{x}).$$

Como apareceu essa última matriz? Não se preocupe, logo estudaremos uma técnica para encontrar esse tipo de matrizes, isto é, matrizes que representam transformações lineares. E, a seguir, veremos que toda transformação linear tem uma matriz associada e vice-versa.

Em geral, se A é uma matriz $m \times n$, podemos definir um *Operador linear* $T:\mathbb{R}^n \to \mathbb{R}^m$ por $T(\mathbf{x}) = A\mathbf{x}$, $\mathbf{x} \in \mathbb{R}^n$.

De fato,

$$\begin{aligned} T(\lambda\mathbf{x}+\mu\mathbf{y}) &= A(\lambda x+\mu y) \\ &= \lambda A\mathbf{x}+\mu A\mathbf{y} \\ &= \lambda T(\mathbf{x})+\mu T(\mathbf{y}) \end{aligned}$$

Podemos, então, considerar cada matriz $A_{m\times n}$ como um operador linear de $\mathbb{R}^n$ em $\mathbb{R}^m$.

Vimos no exemplo anterior que o operador poderia ter sido definido em termos de uma matriz.

Voltemos a reconsiderar as transformações no plano mencionadas anteriormente, agora como sendo definidas através de matrizes:

- Expansão

$$T(\mathbf{x}) = 2x = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} \text{ ou } T(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

- Projeção

$$T(\mathbf{x}) = x_1 e_1 = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \text{ ou } T(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

- Reflexões

$$T(\mathbf{x}) = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} \text{ ou } T(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\text{e } T(\mathbf{x}) = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix} \text{ ou } T(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

- Rotações

$$T(\mathbf{x}) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} \text{ ou } T(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ e}$$

$$T(\mathbf{x}) = \begin{pmatrix} x_1\cos\theta - x_2\operatorname{sen}\theta \\ x_1\operatorname{sen}\theta + x_2\cos\theta \end{pmatrix} \text{ ou } T(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

- Cisalhamento Horizontal

$$T(\mathbf{x}) = \begin{pmatrix} x_1 + ax_2 \\ x_2 \end{pmatrix} \text{ ou } T(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Considere a seguinte matriz (sendo k um escalar): $A = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$.

Se $T(\mathbf{x}) = A\mathbf{x}$, então teremos:

- uma dilatação se $k > 1$;
- a identidade se $k = 1$;
- uma contração se $0 < k < 1$;
- a transformação nula se $k = 0$.

Se V é um espaço vetorial, então o operador identidade é definido por $T(\mathbf{x}) = I\mathbf{x} = \mathbf{x}$. (onde I é a matriz identidade). E, desse modo, o operador leva V em si mesmo.

Reconhecer que uma TL pode ser definida por meio de uma matriz nos permitirá resolver facilmente várias questões práticas e teóricas.

Consideremos uma matriz $A_{m \times n}$ e a transformação definida como $T(\mathbf{x}) = A\mathbf{x}, \mathbf{x} \in \mathbb{R}^n$. Se $\mathbf{u}$ e $\mathbf{y}$ são vetores de $\mathbb{R}^n$ e λ um escalar, temos que:

$$\begin{aligned} T(\lambda\mathbf{u} + \mathbf{y}) &= A(\lambda\mathbf{u} + \mathbf{y}) \\ &= A(\lambda\mathbf{u}) + A\mathbf{y} \\ &= \lambda A\mathbf{u} + A\mathbf{y} \\ &= \lambda T(\mathbf{u}) + T(\mathbf{y}). \end{aligned}$$

E assim provamos que $T(\mathbf{x}) = A\mathbf{x}$ é uma TL.

Para obter essas matrizes foi utilizado um procedimento heurístico, surgido da experiência adquirida quando realizamos produto de matrizes, o que foi possível devido ao fatp de os espaços e as matrizes resultantes serem de dimensão pequena.

Exercícios Propostos

1) Mostre que cada uma das aplicações seguintes é uma transformação linear de $\mathbb{R}^2$ em $\mathbb{R}^2$. Descreva geometricamente o que cada uma delas faz.

a) $T(\mathbf{x}) = (-x_1, x_2)^T$

b) $T(\mathbf{x}) = -x_1$

c) $T(\mathbf{x}) = (x_2, x_1)^T$

d) $T(\mathbf{x}) = \dfrac{1}{2}x$

e) $T(\mathbf{x}) = x_2\mathbf{e}_2$

2) Seja L a transformação linear de $\mathbb{R}^2$ em si mesmo definida por $T(\mathbf{x}) = (x_1 \cos\alpha - x_2 \operatorname{sen}\alpha, x_1 \operatorname{sen}\alpha + x_2 \cos\alpha)^T$. Expresse x_1, x_2 e $T(\mathbf{x})$ em coordenadas polares. Descreva geometricamente o efeito dessa transformação linear.

3) Seja a um vetor fixo não-nulo em $\mathbb{R}^2$. Uma aplicação da forma $T(\mathbf{x}) = x + a$ é chamada de *translação*. Mostre que uma translação não é uma transformação linear. Ilustre geometricamente o efeito de uma translação.

4) Determine se as transformações de $\mathbb{R}^3$ em $\mathbb{R}^2$ a seguir são ou não lineares.

a) $T(\mathbf{x}) = (x_2, x_3)^T$

b) $T(\mathbf{x}) = (0, 0)^T$

c) $T(\mathbf{x}) = (1 + x_1, x_2)^T$

d) $T(\mathbf{x}) = (x_3, x_1 + x_2)^T$

5) Determine se as transformações de $\mathbb{R}^2$ em $\mathbb{R}^3$ a seguir são ou não lineares.

a) $T(\mathbf{x}) = (x_1, x_2, 1)^T$

b) $T(\mathbf{x}) = (x_1, x_2, x_1 + 2x_2)^T$

c) $T(\mathbf{x}) = (x_1, 0, 0)^T$

d) $T(\mathbf{x}) = (x_1, x_2, x_1^2 + x_2^2)^T$

6) Determine se as transformações de $\mathbb{R}^{n\times n}$ em $\mathbb{R}^{n\times n}$ a seguir são ou não lineares.

a) $T(A) = 2A$

b) $T(A) = A^T$

c) $T(A) = A + 1$

d) $T(A) = A - A^T$

7) Determine se as transformações de P_2 em P_3 a seguir são ou não lineares.

a) $T(p(x)) = xp(x)$

b) $T(p(x)) = x^2 + px$

c) $T(p(x)) = p(x) + xp(x) + x^2 p'(x)$

8) Seja $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ uma base para um espaço vetorial V e sejam T_1 e T_2 duas transformações lineares de V em um espaço vetorial W. Mostre que, se $T_1(\mathbf{v}_i) = T_2(\mathbf{v}_i)$ para cada $i = 1, 2, \ldots, n$, então $T_1 = T_2$ [isto é, mostre que $T_1(\mathbf{v}) = T_2(\mathbf{v})$ para todo $\mathbf{v} \in V$].

9) Seja T uma transformação linear de $\mathbb{R}^1$ em $\mathbb{R}^2$ e seja $\mathbf{a} = T(1)$. Mostre que $T(x) = \mathbf{a}x$ para todo $x \in \mathbb{R}^1$, $\mathbf{a} \in \mathbb{R}^2$.

4.2 Operações com Transformações Lineares

4.2.1 Soma, Multiplicação por Escalar, Composição e Inversa

Se aceitarmos que há uma correspondência entre matrizes e transformações lineares, é possível indagar sobre as correspondências entre as operações indicadas nas colunas da seguinte tabela.

Matriz	Transformação linear
Multiplicação por escalar: $c \cdot A$	?
Soma de matrizes: $A + B$	?
Produto de matrizes: $A \cdot B$	?
Inversa de matrizes: A^{-1}	?

As duas primeiras equivalências podem ser respondidas com facilidade. Sabemos que de $A\mathbf{x}$ temos que $(c \cdot A)\mathbf{x} = c(A \cdot \mathbf{x})$ e $(A + B)\mathbf{x} = A\mathbf{x} + B\mathbf{x}$, o que nos induz a introduzir as seguintes definições:

Para a transformação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ e o escalar c, *a multiplicação por escalar* cT está definida como $(cT)(\mathbf{x}) = c(T(\mathbf{x}))$.

E também para as transformações lineares T, $S: \mathbb{R}^n \rightarrow \mathbb{R}^m$, *a adição* $T + S$ está definida pela expressão $(T + S)(\mathbf{x}) = T(\mathbf{x}) + S(\mathbf{x})$.

Exemplo 8. Seja T e S uma reflexão em torno do eixo das abscissas e uma rotação de 90º graus no sentido anti-horário no plano R^2, respectivamente. A figura seguinte ilustra o efeito da soma $T+S$.

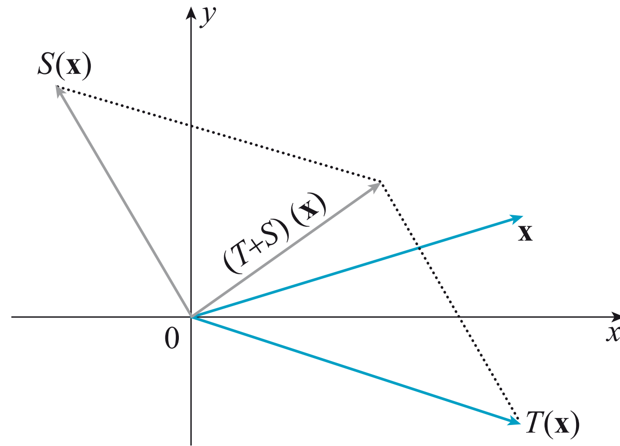


Figura. 4.12 – Soma de transformações

A expressão para a soma é

$$(T+S)(\mathbf{x}) = T(\mathbf{x}) + S(\mathbf{x}) = (x_1, -x_2)^T + (-x_2, x_1)^T = (x_1, -x_2, x_1, -x_2)^T.$$

O que corresponde em termos matriciais a:

$$\left(\overset{\text{reflexão}}{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} + \overset{\text{rotação}}{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}} \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_1 - x_2 \end{pmatrix}$$

Para responder à questão "qual é a operação entre as transformações lineares que corresponde ao produto de matrizes?", teremos de ser um pouco mais cuidadosos.

Se $T:\mathbb{R}^m \to \mathbb{R}^n$ e $S:\mathbb{R}^n \to \mathbb{R}^p$ são duas transformações lineares, podemos aplicar T e depois S para formar a *composta* das duas transformações, que denotamos por $S \circ T$. Note que para que $S \circ T$ faça sentido, o contradomínio de T e o domínio de S devem ser o mesmo (neste caso $\mathbb{R}^n$), e a transformação resultante $S \circ T$ vai do domínio de T ao contradomínio de S, (neste caso $S \circ T:\mathbb{R}^m \to \mathbb{R}^p$). A figura mostra esquematicamente como essa composta funciona. A definição formal da composta de transformações pode ser obtida lembrando da definição correspondente da composta de funções ordinárias.

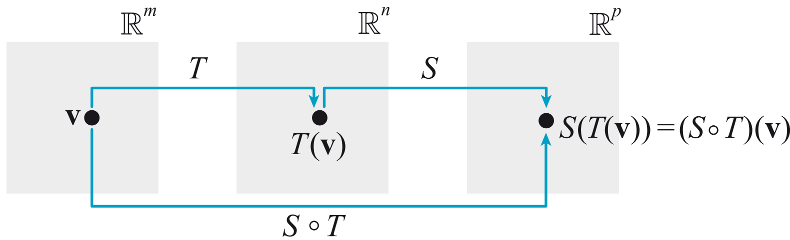


Figura 4.13 – A composta de transformações

Claro que gostaríamos que $S\circ T$ também fosse uma TL. Podemos demonstrar isso verificando que (a composta de transformações) satisfaz a definição de transformação linear (TL) (o que será feito na próxima seção), mas, como no momento assumimos que transformações lineares podem ser definidas por meio de matrizes, usaremos esse fato para prová-lo.

Teorema. *Sejam* $T:\mathbb{R}^m\to\mathbb{R}^n$ *e* $S:\mathbb{R}^n\to\mathbb{R}^p$ *duas transformações lineares e sejam* $A_{p\times n}$ *e* $B_{n\times m}$ *as respectivas matrizes de* S *e* T. *Então* $S\circ T:\mathbb{R}^m\to\mathbb{R}^p$ *é uma transformação linear e suas matrizes são relacionadas por* $C_{p\times m}=A_{p\times n}B_{n\times m}$, *onde* $C_{p\times m}$ *é a matriz de* $S\circ T$.

Demonstração. Se $\mathrm{v}\in\mathbb{R}^m$, aplicando consecutivamente T e S, temos: $T(\mathbf{v})\in\mathbb{R}^n$ e $S(T(\mathbf{v}))=(S\circ T)(\mathbf{v})\in\mathbb{R}^p$.

Assim, temos que:

$$(S\circ T)(\mathbf{v})=S(T(\mathbf{v}))=S(B\mathbf{v})=A(B\mathbf{v})=AB\mathbf{v}=C\mathbf{v}$$

Onde A e B são as respectivas matrizes das transformações S e T.

E então definir uma nova transformação linear, digamos $R:\mathbb{R}^m\to\mathbb{R}^p$ como $R(\mathbf{v})=C\mathbf{v}$.

Agora também podemos dizer que *a matriz da composta é o produto das matrizes e para as transformações lineares* $T:\mathbb{R}^m\to\mathbb{R}^n$ e $S:\mathbb{R}^n\to\mathbb{R}^p$, *a composição de* S *com* T *está definida como* $(S\circ T)(\mathbf{v})=S(T(\mathbf{v}))$.

■

Exemplo 9. Considere novamente o exemplo anterior onde $T(\mathbf{x})=(x_1,-x_2)^T$ e $S(\mathrm{x})=(-x_2,x_1)^T$ e calculemos $S\circ T$ e $T\circ S$.

Solução. $(S\circ T)(\mathbf{x})=S(T(\mathbf{x}))=S(x_1,-x_2)^T=(x_2,x_1)^T$ ou em forma matricial

$$(S \circ T)(\mathbf{x}) = \overset{\text{rotaçaõ}}{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}} \overset{\text{reflexão}}{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$$

Calculemos agora $T \circ S$

$$(T \circ S)(\mathbf{x}) = T(S(\mathbf{x})) = T((-x_2, x_1)^T = (-x_2, -x_1)^T \text{ ou}$$

$$(T \circ S)(\mathbf{x}) = \overset{\text{reflexão}}{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} \overset{\text{rotaçaõ}}{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ -x_1 \end{pmatrix}$$

A definição da composição de transformações lineares nos permitirá introduzir com facilidade a noção de *transformação inversa* e completar a equivalência que fizemos anteriormente.

Considere uma rotação de 90° no sentido horário em torno da origem, seguida por uma rotação de 90° no sentido anti-horário, que denotamos por R_{-90} e R_{90}, respectivamente. Sem dúvida, isso deixa inalterado qualquer vetor de $\mathbb{R}^2$. Podemos expressar esse efeito através de uma composição de R_{90} com R_{-90} como $(R_{90} \circ R_{-90})(\mathbf{v}) = \mathbf{v}$.

Já definimos uma transformação linear que chamamos *identidade* $I(\mathbf{v}) = \mathbf{v}$ de $\mathbb{R}^n$ em $\mathbb{R}^n$, que possui essa característica. Assim, podemos escrever para este caso em particular $I(\mathbf{v}) = (R_{90} \circ R_{-90})(\mathbf{v}) = \mathbf{v}$.

Duas transformações que estão relacionadas desse modo são chamadas de *transformações inversas*.

Definição. Considere T e S como transformações lineares de $\mathbb{R}^n$ em $\mathbb{R}^n$. Então, T e S são transformações inversas se $S \circ T = T \circ S = I$.

Já que nessa definição há uma simetria na relação T *e* S, dizemos que, quando essa situação ocorre, S é a inversa de T e T é a inversa de S. Além disso, dizemos que T e S são *inversíveis*.

Em termos de matrizes, se considerarmos $A_{n \times n}$ e $B_{n \times n}$ as respectivas matrizes de S e T, vemos imediatamente que, se T e S são transformações lineares inversas, tais que $(S \circ T)(\mathbf{x}) = AB\mathbf{x} = I\mathbf{x} = \mathbf{x}$.

Precisamos ter também $(T \circ S)(\mathbf{x}) = BA\mathbf{x} = I\mathbf{x} = \mathbf{x}$ (Verificar!). Isso mostra que A e B são matrizes inversas.

E mostra algo mais: se uma transformação linear T é inversível, sua matriz canônica tem de ser inversível, e já que matrizes inversas são únicas, isso significa que a inversa de T também é única. Portanto, podemos usar a notação de T^{-1} para nos referirmos à inversa de T. Dessa forma, podemos reescrever as equações anteriores como $(T^{-1} \circ T)(\mathbf{x}) = B^{-1}B\mathbf{x} = \mathbf{x} = BB^{-1}\mathbf{x} = (T \circ T^{-1})(\mathbf{x})$, a qual mostra que a matriz de T^{-1} (B^{-1}) é a matriz inversa de $T(B^{-1})$. Acabamos de provar o teorema a seguir.

Teorema. *Se B é uma transformação inversível, então sua matriz canônica também é inversível e $B^{-1} = (B)^{-1}$.*

Lê-se: "a matriz da transformação inversa é a inversa da matriz da transformação original"

Exemplo 10. Determine se a projeção sobre o eixo das abscissas é uma TL inversível.

Solução. A matriz canônica dessa projeção é $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, a qual não é inversível, portanto uma projeção não é inversível. Esse é um fato curioso. Tentemos entender: uma projeção leva ("esmaga") todos os vetores de $\mathbb{R}^2$ sobre o eixo x,

$$T(\mathbf{x}) = x_1 e_1 = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}, \ \forall x \in \mathbb{R}^2.$$

Para "desfazer" esse efeito, ou seja, recuperar $x = (x_1, x_2)^T$ precisaríamos de uma transformação que leve $(x_1, 0)^T$ para $\mathbf{x} = (x_1, x_2)^T$. No entanto, existem infinitos candidatos para serem a imagem de $(x_1, 0)^T$ sob essa hipotética transformação inversa, já que não temos como determinar x_2 no vetor imagem para cada $(x_1, 0)^T$ no domínio de T^{-1}.

$$T(\mathbf{x}) = x_1 e_1 = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \text{ ou } T(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

4.2.2 Transformações Lineares em Espaços de Funções

Nos dois exemplos seguintes são usados conceitos de Cálculo Diferencial e Integral.

$C^1[a,b]$ e $C[a,b]$ são os espaços das funções com primeira derivada contínua e das funções contínuas no intervalo [a,b], respectivamente.

Exemplo 11. Seja o operador derivada de uma função real f, $D : C^1[a,b] \to C[a,b]$, definida por $D(f) = f'$.

D é uma TL, uma vez que

$$\begin{aligned} D(\lambda f + \mu g) &= (\lambda f + \mu g)' \\ &= \lambda f' + \mu g' \\ &= \lambda D(f) + \mu D(g). \end{aligned}$$

Exemplo 12. Seja $T : C[a,b] \to \mathbb{R}$ uma aplicação definida por $T(f) = \int_a^b f(x)\,dx$.

Se f e g são dois vetores em $C[a,b]$, então

$$\begin{aligned} T(\lambda f + \mu g) &= \int_a^b (\lambda f + \mu g)(x)\,dx \\ &= \lambda \int_a^b f(x)\,dx + \mu \int_a^b g(x)\,dx \\ &= \lambda T(f) + \mu T(g). \end{aligned}$$

T é uma TL.

Exemplo 13. Sejam $V = P_n$, $W = P_{n+1}$ espaços de polinômios de grau n e $n+1$, respectivamente, e a aplicação $T : P_n \to P_{n+1}$ definida por:

$$T(p(x)) = xp(x) = a_0 x + a_1 x^2 + \cdots + a_n x^{n+1}.$$

Sejam $p(x)$ e $q(x)$ vetores de P_n e k um escalar, temos que:

$$T(kp) = x(kp)(x) = x(kp(x)) = kxp(x) = kT(p) \quad \text{e}$$

$$T(p+q) = x(p+q)(x) = x(p(x)+q(x)) = xp(x) + xq(x) = T(p) + T(q)$$

Portanto T é uma TL.

Exemplo 14. Seja $V = C(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \ / \ f \text{ é contínua}\}$. Considere a aplicação $J : C(\mathbb{R}) \to \mathbb{R}$, onde $J(f) = f(0)$. Com k um escalar, temos:

$$J(f+g) = (f+g)(0) = f(0) + g(0) = J(f) + J(g) \ \text{ e}$$

$$J(kf) = (kf)(0) = kf(0) = kJ(f)$$

e assim J é uma TL.

Exemplo 15. Seja $V = W = P_n$, $T : P_n \to P_n$ e considere a transformação

$$T(p(x)) = p(ax+b) = a_0 + a_1(ax+b) + a_2(ax+b)^2 + \cdots + a_n(ax+b)^n .$$

Agora temos

$$T(kp) = (kp)(ax+b) = k(p(ax+b)) = kT(p) \text{ e}$$

$$T(p+q) = (p+q)(ax+b) = p(ax+b) + q(ax+b) = T(p) + T(q)$$

Portanto, T é uma TL.

Observação. Note que para provar que estas transformações são lineares foram utilizadas propriedades das funções contínuas e deriváveis.

Nas próximas seções serão estabelecidos teoremas, propriedades e alguns resultados que darão uma estrutura muito útil sobre as transformações lineares.

4.2.3 Propriedades das Transformações Lineares

Se T é uma TL de um espaço vetorial V e um espaço vetorial W, isto é, $T : V \to W$, temos que:

i) $T(\mathbf{0}_v) = \mathbf{0}_w$ (onde $\mathbf{0}_v$ e $\mathbf{0}_w$ são os vetores nulos de V e W, respectivamente)

Prova. Esta afirmação segue-se da condição $T(\lambda \mathbf{u}) = \lambda T(\mathbf{u})$ com $\lambda = 0$. Podemos resumir esta propriedade desta maneira: toda transformação linear leva o vetor nulo no vetor nulo.

Com esta propriedade fica mais fácil verificar que uma translação não é uma TL. Se T(0) ≠ 0, T não é uma TL. Mas, cuidado! T(0) = 0 não é suficiente para que T seja uma TL. Encontre um contra-exemplo!

ii) $T(\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \ldots + \lambda_n \mathbf{u}_n) = \lambda_1 T(\mathbf{u}_1) + \lambda_2 T(\mathbf{u}_2) + \ldots + \lambda_n T(\mathbf{u}_n)$.

Prova. Esta afirmação pode ser provada facilmente usando indução matemática.

Veja que esta igualdade é óbvia para $k = 1$, já que $T(\lambda_1 \mathbf{u}_1) = \lambda_1 T(\mathbf{u}_1)$ porque T é uma TL. Suponha agora que é verdadeira para k:

$$T(\lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \ldots + \lambda_k\mathbf{u}_k) = \lambda_1 T(\mathbf{u}_1) + \lambda_2 T(\mathbf{u}_2) + \ldots + \lambda_k T(\mathbf{u}_k)\,.$$

E agora queremos provar para $k+1$. Somando $\lambda_{k+1}T(\mathbf{u}_{k+1})$ a ambos da última igualdade, obtemos:

$$T(\lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \ldots + \lambda_k\mathbf{u}_k) + \lambda_{k+1}T(\mathbf{u}_{k+1}) = \lambda_1 T(\mathbf{u}_1) + \lambda_2 T(\mathbf{u}_2) + \ldots + \lambda_k T(\mathbf{u}_k) + \lambda_{k+1}T(\mathbf{u}_{k+1}).$$

O lado direito já está da forma que queremos. Usando novamente as condições de TL no termo da esquerda conseguimos escrever esta igualdade da forma:

$$T(\lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \ldots + \lambda_k\mathbf{u}_k + \lambda_{k+1}\mathbf{u}_{k+1}) = \lambda_1 T(\mathbf{u}_1) + \lambda_2 T(\mathbf{u}_2) + \ldots + \lambda_k T(\mathbf{u}_k) + \lambda_{k+1}T(\mathbf{u}_{k+1})$$

que era o que queríamos provar. *A transformação de uma combinação linear de vetores é igual à combinação linear dos vetores transformados.*

Exemplo 16. Determinar qual é a transformação linear T de $\mathbb{R}^2 \to \mathbb{R}^3$, tal que $T = (\mathbf{e}_1) = (2,-1,0)^T$ e $T = (\mathbf{e}_2) = (0,0,1)^T$.

Seja $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$ um vetor arbitrário. Pela aplicação da transformação que estamos procurando sobre este vetor e usando a propriedade anterior temos:

$$\begin{aligned} T(x) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) \\ &= x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) \\ &= x_1(2,-1,0)^T + x_2(0,0,1)^T \\ &= (2x_1, -x_1, x_2)^T. \end{aligned}$$

Exemplo 17. Sejam $V = M_{2\times 2}$ (espaço das matrizes quadradas 2×2) e $W = P_4$ (espaço dos polinômios de grau 4). Qual é a transformação linear $T : M_{2\times 2} \to P_4$, tal que:

$$T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = x^4 + x$$

$$T\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = x^3 + x^2$$

$$T\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = x^2 + x^3$$

$$T\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) = x + x^4$$

Uma matriz $A \in M_{2\times 2}$ é da forma $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ e pode ser escrita como

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

portanto:

$$\begin{aligned} T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) &= T\left(a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) \\ &= T\left(a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) + T\left(b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) + T\left(c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) + T\left(d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) \\ &= aT\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) + bT\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) + cT\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) + dT\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) \\ &= a(x^4 + x) + b(x^3 + x^2) + c(x^2 + x^3) + d(x + x^4) \\ &= (a+d)x + (b+c)x^2 + (b+c)x^3 + (a+d)x^4. \end{aligned}$$

Logo, para todo vetor $\mathrm{v} \in M_{2\times 2}$ (este vetor representa uma matriz), a nossa transformação fica definida como

$$T(\mathrm{v}) = (a+d)x + (b+c)x^2 + (b+c)x^3 + (a+d)x^4,$$

um polinômio de grau 4.

iii) $T(-\mathbf{u}) = -T(\mathbf{u})$

Prova. Para provar esta propriedade observe que:

$$\mathbf{0}_w = T(\mathbf{0}_v) = T(\mathbf{u} + (-\mathbf{u})) = T(\mathbf{u}) + T(-\mathbf{u})$$

O que significa que $T(-\mathbf{u})$ é o inverso aditivo de $T(\mathbf{u})$, isto é, $T(-\mathbf{u}) = -T(\mathbf{u})$.

Exemplo 18. Uma transformação linear importante é aquela que se obtém usando-se o produto escalar. Seja o espaço $\mathbb{R}^n$ com o produto escalar usual $\langle .,. \rangle$ e $\mathbf{v}_0 \in \mathbb{R}^n$ um vetor qualquer prefixado. Seja a aplicação $T : \mathbb{R}^n \to \mathbb{R}$ definida por $T(\mathbf{v}) = \langle \mathbf{v}_0, \mathbf{v} \rangle$. Mostraremos que é uma TL utilizando as propriedades do produto escalar. Verifiquemos as condições das transformações lineares. Sejam $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ e k um escalar, temos que:

$$T(k\mathbf{v}) = \langle \mathbf{v}_0, k\mathbf{v} \rangle = k\langle \mathbf{v}_0, \mathbf{v} \rangle = kT(\mathbf{v}) \text{ e}$$

$$T(\mathbf{v}+\mathbf{w}) = \langle \mathbf{v}_0, (\mathbf{v}+\mathbf{w}) \rangle = \langle \mathbf{v}_0, \mathbf{v} \rangle + \langle \mathbf{v}_0, \mathbf{w} \rangle = T(\mathbf{v}) + T(\mathbf{w})$$

assim, o produto escalar é uma TL.

4.2.4 Composição de Transformações Lineares

Em uma seção anterior, definimos a composição de transformações lineares por matrizes. A definição se estende de maneira natural para TL em geral.

Definição. Se $T: U \to V$ e $S: V \to W$ são transformações lineares, a composição de S com T é a aplicação que denotamos por $S \circ T$ (lê-se "S composta com T") definida por $(S \circ T)(\mathrm{u}) = S(T(\mathrm{u}))$, onde u é um elemento de U. Observe que $S \circ T$ é uma aplicação de U em W. Note também que, para a definição fazer sentido, a imagem de T deve estar contida no domínio de S.

Exemplo 19. Sejam duas transformações lineares definidas por:

$$T: \mathbb{R}^2 \to P_1, \quad T(\mathbf{a}) = a_1 + (a_1 + a_2)x$$
$$S: P_1 \to P_2, \quad S(p(x)) = xp(x)$$

Encontre $(S \circ T)(\mathrm{a})$.

Solução.

$$(S \circ T)(\mathbf{a}) = S(T(\mathbf{a}) = S(a_1 + (a_1 + a_2)x) = x(a_1 + (a_1 + a_2)x) = a_1x + (a_1 + a_2)x^2$$

Teorema. *Se $T: U \to V$ e $S: V \to W$ são transformações lineares, então $T: U \to W$ é uma transformação linear.*

Demonstração. Sejam $\mathbf{u}$ e $\mathbf{v}$ em U e $\lambda \in \mathbb{R}$ um escalar, iremos conferir as condições de TL usando o fato de que S e T são transformações lineares:

$$\begin{aligned}(S \circ T)(\mathbf{u}+\mathbf{v}) &= S(T(\mathbf{u}+\mathbf{v})) \\ &= S(T(\mathbf{u}) + T(\mathbf{v})) \\ &= S(T(\mathbf{u})) + S(T(\mathbf{v})) \\ &= (S \circ T)(\mathbf{u}) + (S \circ T)(\mathbf{v})\end{aligned}$$

$$\begin{aligned}(S \circ T)(\lambda\mathbf{v}) &= S(T(\lambda\mathbf{v})) \\ &= S(\lambda T(\mathbf{v})) \\ &= \lambda S(T(\mathbf{v})) \\ &= \lambda(S \circ T)(\mathbf{v})\end{aligned}$$

■

As propriedades algébricas das transformações lineares espelham as das transformações matriciais, as quais, por sua vez, estão relacionadas com as propriedades algébricas das matrizes. Por exemplo, a composição das transformações lineares é associativa. Para ver isto, considere as seguintes transformações lineares R, S e T, então $R \circ (S \circ T) = (R \circ S) \circ T$, desde que essas composições façam sentido.

4.3 A Imagem e o Núcleo de uma Transformação Linear

Definição. Seja $T: V \to W$ uma transformação linear e seja S um subespaço de V ($S \subset V$). A *imagem* de S, denotada por $T(S)$ é o conjunto de vetores $\mathbf{w} \in W$ tais que existe um vetor $\mathbf{v} \in W$, que satisfaça $T(\mathbf{v}) = \mathbf{w}$. Ou seja

$$\text{Im}(T) = T(S) = \{\mathbf{w} \in W \,/\, T(\mathbf{v}) = \mathbf{w} \text{ para algum } \mathbf{v} \in S\} \subseteq W.$$

A imagem de todo o espaço vetorial V (quando $S = V$), $\text{Im}(T(V))$, é chamada de *imagem* de T.

A palavra *kernel* é derivada do termo *cyrnel*, do inglês antigo, uma forma da palavra *corn*, que significa grão. Tal como um grão, o *kernel* de uma transformação linear é o seu núcleo, no sentido de carregar muitas informações sobre propriedades importantes da transformação.

Definição. Seja $T: V \to W$ uma transformação linear. O conjunto de todos os vetores $\mathbf{v} \in V$ em que $T(\mathbf{v}) = 0$ é chamado *núcleo* de T, sendo denotado por $\text{Ker}(T)$. Isto é,

$$\text{Ker}(T) = \{\mathbf{v} \in V \,/\, T(\mathbf{v}) = \mathbf{0}_w\} \subseteq V.$$

É fácil ver que $\text{Ker}(T)$ é um subconjunto de V e que, se S é um subespaço qualquer de V, então $\text{Im}(T(S))$ é um subconjunto de W.

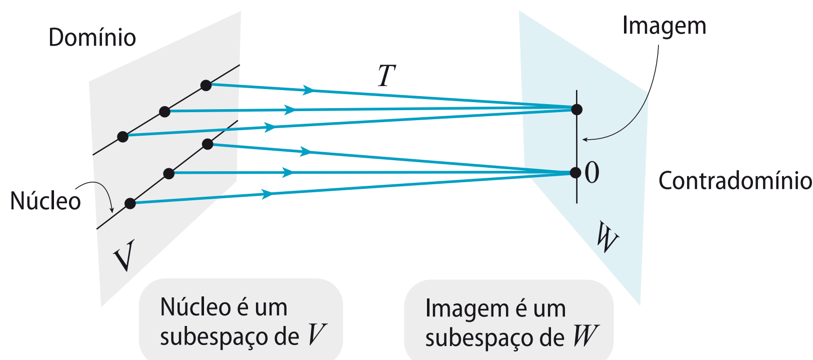


Figura 4.14 – Subespaços associados a uma transformação linear

Teorema. *Se $T: V \rightarrow W$ é uma transformação linear e S é um subespaço de V, então:*

- $\text{Ker}(T)$ *é um subespaço de V*

  Demonstração. Vamos mostrar que o núcleo de V é fechado sob as operações de adição e multiplicação por escalar.

  Seja $\text{v} \in \text{Ker}(T)$, isto é, $T(\mathbf{v}) = \mathbf{0}_w$, e α é um escalar, então:

  $$T(\alpha\mathbf{v}) = \alpha T(\mathbf{v}) = a\mathbf{0}_w = \mathbf{0}_w \Rightarrow \alpha\mathbf{v} \in \text{Ker}(T).$$

  Por outro lado, se $\mathbf{v}_1, \mathbf{v}_2 \in \text{Ker}(T)$, isto é, $T(\mathbf{v}_1) = \mathbf{0}_w$ e $T(\mathbf{v}_2) = \mathbf{0}_w$, então:

  $$T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) = \mathbf{0}_w + \mathbf{0}_w = \mathbf{0}_w \Rightarrow \mathbf{v}_1 + \mathbf{v}_2 \in \text{Ker}(T).$$

  ■

- $\text{Im}(T(S))$ *é um subespaço de W*

  Demonstração. A demonstração é semelhante à anterior. Se $\mathbf{w} \in \text{Im}(T(S))$, então $\mathbf{w} = T(\mathbf{v})$ para algum $\mathbf{v} \in S$.

  Para qualquer escalar α, temos $\alpha\mathbf{w} = \alpha T(\mathbf{v}) = T(\alpha\mathbf{v})$. Como $\alpha\mathbf{v} = S$, $a\mathbf{w} = \text{Im}(T(S))$ e $T(S)$ é fechado sob a multiplicação por escalar.

  Se $\mathbf{w}_1, \mathbf{w}_2 \in \text{Im}(T(S))$, existem $\mathbf{v}_1, \mathbf{v}_2 \in V$ tais que $T(\mathbf{v}_1) = \mathbf{w}_1$ e $T(\mathbf{v}_2) = \mathbf{w}_2$. Logo, $\mathbf{w}_1 + \mathbf{w}_2 = T(\mathbf{v}_1) + T(\mathbf{v}_2) = T(\mathbf{v}_1 + \mathbf{v}_2)$ e $\text{Im}(T(S))$ é fechado sob a adição.

  ■

Exemplo 20. Seja a transformação $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $T(\mathbf{x}) = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$.

Um vetor $\mathbf{x}$ pertence ao núcleo de T se e somente se $x_1 = 0$, isto é,

$$T(\mathbf{x}) = 0 \Leftrightarrow \begin{pmatrix} x_1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow x_1 = 0.$$

Logo, os vetores de $\text{Ker}(T)$ têm primeira componente nula, $\mathbf{x} = \begin{pmatrix} 0 \\ x_2 \end{pmatrix} = 0\mathbf{e}_1 + x_2\mathbf{e}_2 = x_2\mathbf{e}_2$, então $\text{Ker}(T)$ é o subespaço unidimensional de $\mathbb{R}^2$ gerado por $\mathbf{e}_2$.

Escrevendo $y = T(\mathbf{x})$, o vetor $\mathbf{y}$ estará na imagem de T se e somente se $\mathbf{y}$ é um múltiplo de x_1, isto é,

$$\mathbf{y} = T(\mathbf{x}) = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \Leftrightarrow \mathbf{y} = x_1\mathrm{e}_1 + 0\mathrm{e}_2 = x_1\mathrm{e}_1.$$

Logo, $\mathrm{Im}(T) = T(\mathbb{R}^2) = \mathrm{e}_1$ é o subespaço unidimensional de $\mathbb{R}^2$ gerado por $\mathbf{e}_1$.

Exemplo 21. Seja $T:\mathbb{R}^3 \to \mathbb{R}^2$ a transformação linear definida por $T(\mathbf{x}) = (x_1 + x_2, x_2 + x_3)^T$.

O núcleo de T pode ser determinado impondo a condição

$$T(\mathbf{x}) = 0 \Leftrightarrow (x_1 + x_2, x_2 + x_3)^T = (0,0)^T$$

ou equivalentemente, $\mathbf{x} \in \mathrm{Ker}(T)$ se e somente se

$$(x_1 + x_2, x_2 + x_3)^T = (0,0)^T \Leftrightarrow \begin{cases} x_1 + x_2 = 0 \\ x_2 + x_3 = 0 \end{cases}.$$

A solução deste sistema está dada (fazendo a variável livre $x_3 = a$) por todos os vetores de $\mathbb{R}^3$ da forma $a(1,-1,1)^T$.

Logo, $\mathrm{Ker}(T) = \{\mathbf{x} \in \mathbb{R}^3 / \mathbf{x} = a(1,-1,1)^T, a \in \mathbb{R}\}$.

Agora passemos a determinar a imagem, que deve ser um subespaço de $\mathbb{R}^2$. Se $\mathrm{y} \in \mathrm{Im}(T)$, então $\mathbf{y} = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \end{pmatrix} = x_1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2\begin{pmatrix} 1 \\ 1 \end{pmatrix} + x_3\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Então os vetores da imagem são gerados pela combinação linear dos vetores $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$, que geram todo $\mathbb{R}^2$, logo $\mathrm{Im}(T) = \mathbb{R}^2$.

Exemplo 22. Seja $T:\mathbb{R}^3 \to \mathbb{R}^3$ a projeção ortogonal sobre o plano x_1x_2, $T(\mathbf{x}) = (x_1, x_2, 0)^T$.

Temos que $(x_1, x_2, 0)^T = (0,0,0)^T \Leftrightarrow x_1 = 0, x_2 = 0$. Como nada é dito sobre a variável x_3, temos que x_3 é qualquer; logo, $\mathrm{Ker}(T) = \{\mathbf{x} \in \mathbb{R}^3 / \mathbf{x} = a(0,0,1)^T, a \in \mathbb{R}\}$, ou seja, o núcleo de T são todos os vetores que estão sobre o eixo x_3.

A imagem desta transformação é dada diretamente pela sua definição, o plano x_1x_2.

Para determinar o núcleo de uma transformação linear $T:\mathbb{R}^n \to \mathbb{R}^n$ será necessário resolver um sistema homogêneo.

Exercício. Mostre que as transformações lineares de $\mathbb{R}^3$ em $\mathbb{R}^3$, $T(\mathbf{x}) = (0,\ x_2,\ x_3)^T$ e $S(\mathbf{x}) = (0,\ x_2 + x_3,\ 2x_2 + x_3)^T$ têm os mesmos núcleos e imagens.

Resolução. O núcleo da transformação de T é dado pelo conjunto: $\ker(T) = \{\mathbf{x} \in \mathbb{R}^3 : T(\mathbf{x}) = 0\}$; que equivale a escrever $(0, x_2, x_3)^T = (0,0,0)^T$ assim, determinar o núcleo consiste em resolver o sistema $x_2 = 0,\ x_3 = 0$ cuja solução é da forma $(x_1, x_2, x_3)^T = x_1(1,0,0)^T$.

Analogamente, o núcleo da transformação de S é dado pelo conjunto: $\ker(S) = \{\mathbf{x} \in \mathbb{R}^3 : S(\mathbf{x}) = 0\}$; ou seja $(0, x_2 + x_3, 2x_2 + x_3)^T = (0,0,0)^T$ que nos conduz ao sistema $x_2 + x_3 = 0,\ 2x_2 + x_3 = 0$ cuja solução também são os vetores $(x_1, x_2, x_3)^T = x_1(1,0,0)^T$.

Logo $\ker(T) = \ker(S)$.

O contradomínio, ou imagem de T, denotado por $Im(T)$ ou $T(\mathbb{R}^3)$ está dado pelo conjunto $Im(T) = \{\mathbf{y} \in R^3 : \mathbf{y} = T(\mathbf{x}), \forall\, \mathbf{x} \in \mathbb{R}^3\}$, ou seja $(y_1, y_2, y_3)^T = (0, x_2, x_3)^T$.

Assim os vetores $\mathbf{y}$ estão dados por

$$(y_1, y_2, y_3)^T = (0, x_2, x_3)^T = x_2(0,1,0)^T + x_3(0,0,1)^T$$

e então $Im(T) = \{(0,1,0)^T, (0,0,1)^T\}$.

A $Im(S)$ está definida como:

$$Im(S) = \{\mathbf{y} \in \mathbb{R}^3 : \mathbf{y} = T(\mathbf{x}), \forall\, \mathbf{x} \in \mathbb{R}^3\} \text{ ou}$$

$$(y_1, y_2, y_3)^T = (0, x_2 + x_3,\ 2x_2 + x_3)^T$$

que nos leva a

$$(y_1, y_2, y_3)^T = (0, x_2 + x_3,\ 2x_2 + x_3)^T = x_2(0,1,2)^T + x_3(0,1,1)^T.$$

Mostrar que $T(\mathbf{x})$ e $S(\mathbf{x})$ possuem a mesma imagem é equivalente a mostrar que os espaços gerados pelas respectivas imagens são iguais.

Para isso é suficiente mostrar que um mesmo vetor $\mathbf{y}$ nestes espaços pode ser escrito como combinação linear dos vetores em ambas bases: $Im(T)$ e $Im(S)$.

Se $\mathbf{y} \in [(0,1,0)^T, (0,0,1)^T]$ e $\mathbf{y} \in [(0,1,2)^T, (0,1,1)^T]$ então

$$(y_1, y_2, y_3)^T = a_1(0,1,0)^T + a_2(0,0,1)^T = b_1(0,1,2)^T + b_2(0,1,1)^T,$$

usando a segunda igualdade obtemos o seguinte sistema:

$$a_1 = b_1 + b_2,\ a_2 = 2b_1 + b_2$$

cuja solução nos permite escrever

$$(y_1, y_2, y_3)^T = a_1(0,1,0)^T + a_2(0,0,1)^T = (a_2 - a_1)(0,1,2)^T + (2a_1 - a_2)(0,1,1)^T$$

e assim mostramos que as imagens de ambas transformações são iguais.

Exemplo 23. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ a transformação linear que é a projeção ortogonal sobre a reta cujas equações paramétricas são:

$$x_1 = 1 + 2t$$
$$x_2 = 2 - 2t$$
$$x_3 = 3 + t.$$

Ache a transformação linear, seu núcleo e imagem.

Solução. Estas são as equações paramétricas de uma reta de $\mathbb{R}^3$ que passa pelo ponto $(1,2,3)$ na direção do vetor $\mathrm{u} = (2,-2,1)^T$ (o vetor diretor da reta). Projetar um vetor sobre uma reta é o mesmo que encontrar a projeção ortogonal sobre o vetor diretor dessa mesma reta.

$$T(\mathbf{x}) = proj_u \mathbf{x} = \langle \mathbf{x}, \mathbf{u} \rangle \hat{\mathbf{u}} \text{ onde } \hat{\mathbf{u}} = \frac{\mathbf{u}}{|\mathbf{u}|}.$$

Para o nosso caso

$$T(\mathbf{x}) = \left\langle (x_1, x_2, x_3)^T, (2,-2,1)^T \right\rangle \frac{(2,-2,1)^T}{9} =$$

$$= \frac{1}{9}(4x_1 - 4x_2 + 2x_3, -4x_1 + 4x_2 - 2x_3, 2x_1 - 2x_2 + x_3)^T = (0,0,0)^T.$$

Para encontrar o núcleo devemos ter,

$$\frac{1}{9}(4x_1 - 4x_2 + 2x_3, -4x_1 + 4x_2 - 2x_3, 2x_1 - 2x_2 + x_3)^T = (0,0,0)^T.$$

Que leva ao seguinte sistema homogêneo

$$\begin{aligned} 4x_1 - 4x_2 + 2x_3 &= 0 \\ -4x_1 + 4x_2 - 2x_3 &= 0 \\ 2x_1 - 2x_2 + x_3)^T &= 0 \end{aligned}$$

cuja solução é $(a, b. - 2a - 2b)^T$.

Portanto,

$$\text{Ker}(T) = \{\mathbf{x} \in \mathbb{R}^3 \,/\, \mathbf{x} = (x_1, x_2, -2x_1 - 2x_2)^T\} = [(1,0,-2)^T, (0,1-2)^T].$$

Para determinar a imagem, fazemos

$$\begin{pmatrix} 4x_1 - 4x_2 + 2x_3 \\ -4x_1 + 4x_2 - 2x_3 \\ 2x_1 - 2x_2 + x_3)^T \end{pmatrix} = x_1 \begin{pmatrix} 4 \\ -4 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} -4 \\ 4 \\ -2 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}.$$

Logo, todo vetor que pertence à imagem de T é gerado pelos vetores

$$y_1 = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix},\ y_2 = \begin{pmatrix} -2 \\ 2 \\ -1 \end{pmatrix},\ y_3 = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}.$$

que são linearmente dependentes. Podemos então escrever que $\text{Im}(T) = [(2,-2,1)^T]$ (Verificar o porquê!) e é um subespaço unidimensional de $\mathbb{R}^3$. Veja que este vetor coincide com o vetor diretor da reta como era de se esperar.

Exercícios

12) Determine o núcleo e a imagem de cada uma das transformações lineares de $\mathbb{R}^3$ em $\mathbb{R}^3$.

a) $T(\mathbf{x}) = (x_3, x_2, x_1)^T$

b) $T(\mathbf{x}) = (x_1, x_2, 0)^T$

c) $T(\mathbf{x}) = (x_1, x_1, x_1)^T$

13) Seja S o subespaço de $\mathbb{R}^3$ gerado por $\mathbf{e}_1$ e $\mathbf{e}_2$. Para cada um dos operadores lineares no Exercício 12, determine $L(S)$.

14) Determine o núcleo e a imagem de cada uma das transformações lineares de P_3 em P_3 dadas a seguir.

a) $T(p(x)) = xp'(x)$

b) $T(p(x)) = p(x) - p'(x)$

c) $T(p(x)) = p(x) + p(1)$

15) Seja $L: V \to W$ uma transformação linear e seja T um subespaço de W. A *imagem inversa* de T, denotada por $L^{-1}(T)$, é definida por $L^{-1}(T) = \{\mathbf{v} \in V \mid L(\mathbf{v}) \in T\}$. Mostre que $L^{-1}(T)$ é um subespaço de W.

16) Uma transformação linear $T: V \to W$ é dita *injetora* se $T(v_1) = T(v_2)$ implica que $\mathbf{v}_1 = \mathbf{v}_2$ (isto é, dois vetores distintos $\mathbf{v}_1, \mathbf{v}_2 \in V$ não podem ser levados no mesmo vetor $\mathbf{w} \in W$). Mostre que T é injetora se e somente se $\text{Ker}(T) = \{\mathbf{0}_v\}$.

17) Um operador linear $T: V \to W$ é dito *sobrejetora* se $T(V) = W$. Mostre que o operador $T: \mathbb{R}^3 \to \mathbb{R}^3$ definido por $T(\mathbf{x}) = (x_1, x_1 + x_2, x_1 + x_2 + x_3)^T$ é sobrejetora.

18) Quais dos operadores no Exercício 15 são injetores? Quais são sobrejetores?

19) Seja A uma matriz 2×2 e seja T_A o operador definido por $T_A = A\mathbf{x}$. Mostre que:

a) T_A leva $\mathbb{R}^2$ no espaço coluna de A.

b) Se A é inversível, então T_A é sobrejetora de $\mathbb{R}^2$ em $\mathbb{R}^2$.

20) Seja D o operador derivada em P_3 e seja $S = \{p \in P_3 \mid p(0) = 0\}$.

Mostre que:

a) D de P_3 em P_2 é sobrejetora, mas não é injetora;

b) $D: S \to P_3$ é injetora, mas não e sobrejetora.

4.4 Transformações Injetoras, Sobrejetoras e Isomorfismos

Seria conveniente recordar nesse momento as noções de função injetora e sobrejetora, pois estenderemos estes conceitos para as transformações lineares e posteriormente estabeleceremos a relação entre estes conceitos e os de núcleo e imagem.

Definição. Dada uma aplicação $T:V \to W$, diremos que T é *injetora* se dados $\mathbf{u}, \mathbf{v} \in V$ com $T(\mathbf{u}) = T(\mathbf{v})$ tivermos $\mathbf{u} = \mathbf{v}$. Ou equivalentemente, com $\mathbf{u} \neq \mathbf{v}$, então $T(\mathbf{u}) \neq T(\mathbf{v})$. Em outras palavras, T é injetora se as imagens de dois vetores distintos são distintas.

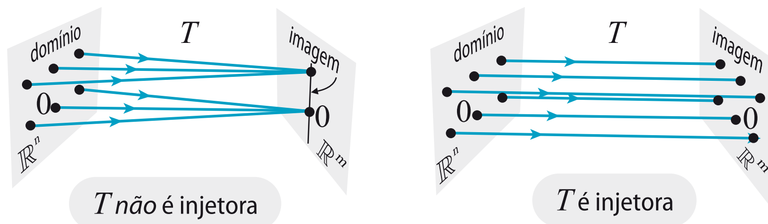


Figura 4.15 – Transformações injetoras e não injetoras

Definição. Uma aplicação $T:V \to W$ será *sobrejetora* se a imagem de T coincidir com W, ou seja, $T(V) = W$.

Observação. Na definição anterior, vimos que uma função será sobrejetora se, dado $\mathbf{w} \in W$ existir $\mathbf{v} \in W$, tal que $T(\mathbf{v}) = \mathbf{w}$.

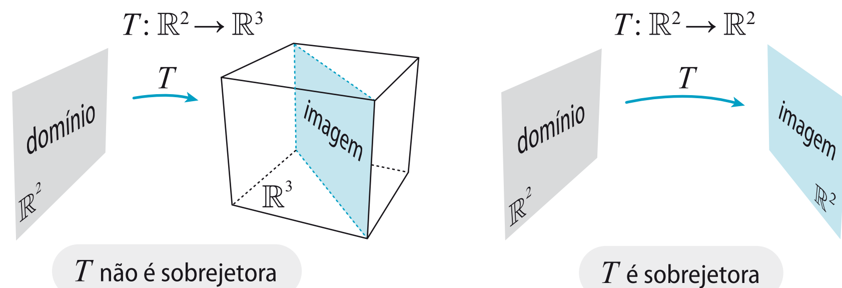


Figura 4.16 – Transformações sobrejetoras e não sobrejetoras

Este teorema afirma que uma TL injetora só tem o vetor nulo no seu núcleo. E, por outro lado, se uma TL tiver somente 0 no seu núcleo, então quaisquer dois vetores distintos devem ter imagens distintas também.

Teorema. *Seja $T:V \to W$, uma aplicação linear. Então,* $\text{Ker}(T) = \{\mathbf{0}\}$ *se e somente se T é injetora.*

$$\text{Ker}(T) = \{0\} \Leftrightarrow T \text{ é injetora}.$$

Corolário. Seja $T:V \to W$ uma aplicação linear injetora. Então, T leva vetores linearmente independentes em vetores linearmente independentes.

Teorema. *Seja $T: V \to W$, uma aplicação linear. Então,*

$$\dim(\ker(T)) + \dim(\text{Im}(T)) = \dim(V).$$

Corolário. Se $\dim V = \dim W$, então T é injetora se T é sobrejetora.

Corolário. Seja $T: V \to W$ uma aplicação linear injetora. Se $\dim V = \dim W$, então T leva base em base.

Exemplo 24. Seja $T: P_n \to P_n$, dada por $T(p(x)) = xp(x)$. Verifique se T é bijetora.

Solução. Devemos verificar se T é injetora e sobrejetora ao mesmo tempo. Para isso, usaremos os teoremas e corolários dados anteriormente.

Para ver que é injetora, devemos apenas calcular o núcleo de T:

$$T(p(x)) = xp(x)$$
$$T(a_0 + a_1x + \cdots + a_nx^n) = x(a_0 + a_1x + \cdots + a_nx^n)$$
$$= a_0x + a_1x^2 + \cdots + a_nx^{n+1}.$$

Quais os polinômios tal que $T(p(x)) = 0$?

$$T(p(x)) = 0 \Leftrightarrow a_0x + a_1x^2 + \cdots + a_nx^{n+1} \Leftrightarrow a_0 = a_1 = \cdots = a_n = 0 \Leftrightarrow p(x) = 0$$

$p(x)$ é o polinômio nulo $\text{Ker}(T) = \{p \in P_n \,/\, p(x) = 0, \forall x \in \mathbb{R}\} = \{0\}$ (observe que neste caso o vetor nulo de P_n é o polinômio nulo de grau n). Portanto, T é injetora.

Como $\dim P_n = n+1$ e $\dim P_{n+1} = n+2$ e $\dim(\text{Ker}(T)) = 0$, temos que:

$$\dim(\ker(T)) + \dim(\text{Im}(T)) = n+1$$
$$0 + \dim(\text{Im}(T)) = n+1$$
$$\dim(\text{Im}(T)) = n+1$$

Note que $\dim(\text{Im}(T)) = n+1 \neq \dim P_{n+1} = n+2$, então: $\text{Im}(T) \neq P_{n+1}$. Portanto, T não é sobrejetora.

Quais são os polinômios que estão faltando na imagem de T?

Quando uma transformação linear $T: V \to W$ for injetora e sobrejetora ao mesmo tempo, dá-se o nome de isomorfismo.

Estes termos são derivados de dois vocábulos gregos: "iso" que significa igual e "morfos" que significa forma.

Sob o ponto de vista da Álgebra Linear, espaços vetoriais isomorfos são, por assim dizer, idênticos.

Temos então que espaços isomorfos devem ter a mesma dimensão. Logo, um isomorfismo leva base em base. Alem disso, um isomorfismo $T:V \to W$ tem uma aplicação inversa, que é linear e é também um isomorfismo.

De forma análoga ao que fizemos com o conceito de composição de transformações lineares, voltamos para as nossas definições das transformações inversas definidas por matrizes. Repetiremos estas definições para as transformações inversas, num contexto mais geral. Estamos agora em condições de estabelecer relações e propriedades com maior formalidade.

Definição. Seja $T:V \to W$ uma transformação linear. Se existe uma aplicação $T^{-1}:V \to W$ tal que $T \circ T^{-1} = I$ e $T^{-1} \circ T = I$, então dizemos que T é inversível e que T^{-1} é a inversa de T.

Propriedades:

1) Seja $T:V \to W$ uma transformação linear. Se existe T^{-1}, a inversa de T, então T^{-1} é uma transformação linear.

2) Se T é um isomomorfismo, então T é inversível e além disso, T^{-1} também é um isomorfismo.

Podemos concluir que, se $T:V \to W$ é uma transformação linear inversível (um isomorfismo), $\mathrm{E} = [\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n]$ e $\mathrm{F} = [\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_m]$ são bases para V e W, A é a matriz de representação para T, então A é inversível.

Dois espaços vetoriais são isomorfos se existe uma transformação linear inversível entre eles. Uma transformação linear com essa propriedade é um isomorfismo.

Em particular, se um espaço vetorial tem uma base de dimensão finita, então é isomorfo com o espaço euclidiano. Lembramos que se um espaço vetorial é gerado por um número finito de vetores, então tem uma base de dimensão finita.

Exemplo 25. O espaço vetorial dos polinômios P_n com coeficientes reais de grau máximo n, é isomorfo com $\mathbb{R}^{n+1}$?

Solução. Para prová-lo, devemos construir um isomorfismo, isto é, uma aplicação linear bijetora $P_n \to \mathbb{R}^{n+1}$. Para isso consideremos as bases canônicas de $\mathbb{R}^{n+1}$ e P_n, $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_{n+1}\}$ e $\{1, t, t^2, \ldots, t^n\}$, respectivamente, e definimos, $T(1) = \mathbf{e}_1, T(t) = \mathbf{e}_2, ..., T(t^n) = \mathbf{e}_{n+1}$. Consideremos a aplicação linear de T em um vetor de P_n e usemos a última definição dada

$$
\begin{aligned}
T(a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n) &= a_0 T(1) + a_1 T(t) + \ldots + a_n T(t^n) \\
&= a_0 \mathbf{e}_1 + a_1 \mathbf{e}_2 + \ldots a_n \mathbf{e}_{n+1} \\
&= (a_0, a_1, a_2 \ldots, a_n)^T.
\end{aligned}
$$

Claramente T é linear (pela própria construção) e $\text{Ker}(T) = \{0\}$. Logo, é injetora.

Como $\dim(\mathbb{R}^{n+1}) = \dim(P_n) = n+1$ é sobrejetora, logo é um isomorfismo e é dado como

$$a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \Leftrightarrow (a_0, a_1, a_2, \ldots a_n)^T.$$

Observação. Este fato pode ser usado na conversão de alguns problemas envolvendo polinômios em vetores de espaços euclidianos.

Exemplo 26. O espaço vetorial das matrizes $M_{2\times 2}$ é isomorfo com $\mathbb{R}^4$. As coordenadas em relação à base

$$\left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right]$$

é o isomorfismo $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \leftrightarrow (a, b, c, d)^T$.

Para ver isto, proceda em forma análoga ao exemplo anterior.

Dado um isomorfismo, resolver uma questão da Álgebra Linear (que pode ser uma definição, um conceito, uma propriedade, um teorema e sua prova, etc.) em um espaço vetorial é equivalente a resolver a mesma questão em um outro espaço vetorial, onde tenhamos alguma familiaridade ou facilidade, ou ainda a questão a ser resolvida seja mais simples. Em particular, para espaços vetoriais finitos, as coordenadas identificam a álgebra linear de um espaço vetorial mais geral com a álgebra linear de um espaço euclidiano.

Para uma transformação linear, $T: V \to W$ as seguintes afirmações são equivalentes:

- T é inversível.
- T é bijetora.
- $\operatorname{Im}(T) = W$, $\ker(T) = \{0\}$.

Atenção: aproveite a informação que pode ser obtida analisando uma matriz para determinar o que quer saber de uma TL.

Se T for uma TL definida por uma matriz A, na forma $T(\mathrm{x}) = A\mathrm{x}$ então o núcleo $\ker(T)$ e a imagem $\operatorname{Im}(T)$ de T podem ser determinados através da análise da matriz A na sua forma reduzida por linhas.

Exemplo 27. Dada a transformação linear $T: \mathbb{R}^3 \to \mathbb{R}^3$, definida por

$$T(\mathbf{x}) = (x_1 + 2x_2 + 2x_3, x_1 + x_2 + 3x_3, x_1 + 2x_2 + x_3)^T$$

a) mostre que existe a inversa de T

b) explicite $T^{-1}(\mathbf{x})$

Solução. Veja que $\dim V = \dim W = 3$ $(V = W = \mathbb{R}^3)$; mostraremos que $\operatorname{Ker}(T) = \{0\}$ e usaremos o corolário adequado para concluir que é um isomorfismo.

Sabemos que $T^{-1}: \mathbb{R}^3 \to \mathbb{R}^3$ é uma TL, tal que $T \circ T^{-1} = T^{-1} \circ T = I$.

Uma forma de proceder é da seguinte maneira: determinamos a matriz da transformação (não deixe de fazer isso, deve ser 3×3) e escrevemos a transformação como $T(\mathrm{x}) = A\mathrm{x} = \mathrm{y}$, logo formamos a matriz aumentada $[A \mid \mathrm{y}]$ e a levamos na forma escalonada reduzida por linhas

$$\begin{pmatrix} 1 & 2 & 2 & y_1 \\ 1 & 1 & 3 & y_2 \\ 1 & 2 & 1 & y_3 \end{pmatrix} \leftrightarrow \begin{pmatrix} 1 & 0 & 0 & -5y_1 + 2y_2 + 4y_3 \\ 0 & 1 & 0 & 2y_1 - y_2 - y_3 \\ 0 & 0 & 1 & y_1 - y_3 \end{pmatrix}.$$

Agora resolveremos praticamente ambos os itens em forma simultânea: para determinar $\operatorname{Ker}(T)$ devemos encontrar a solução de $T(\mathbf{x}) = \mathbf{0}$ para o qual fazemos $\mathbf{y} = \mathbf{0}$; resolvendo o sistema homogêneo temos que $\operatorname{Ker}(T) = \{\mathbf{0}\}$.

A expressão explícita da inversa está dada pela última coluna da forma escalonada reduzida, logo

$$T^{-1}(\mathbf{y}) = (-5y_1 + 2y_2 + 4y_3, 2y_1 - y_2 - y_3, y_1 - y_2)^T.$$

Consideremos uma outra técnica para obter a inversa.

Temos que $(T \circ T^{-1})(\mathbf{x}) = T(T^{-1}(\mathbf{x})) = I(\mathbf{x}) = \mathbf{x}$. Pondo $y = T^{-1}(\mathbf{x})$ temos $T(\mathbf{y}) = \mathbf{x}$, ou seja,

$$T(\mathbf{y}) = (y_1 + 2y_2 + 2y_3, y_1 + y_2 + 3y_3, y_1 + 2y_2 + y_3)^T = (x_1, x_2, x_3)^T.$$

Esta equação pode ser reescrita como

$$\begin{cases} y_1 + 2y_2 + 2y_3 = x_1 \\ y_1 + y_2 + 3y_3 = x_2 \\ y_1 + 2y_2 + y_3 = x_3 \end{cases}$$ e resolvendo para y temos

$$\begin{cases} y_1 = x_1 - x_3 \\ y_2 = 2x_1 - x_2 - x_3 \\ y_3 = -5x_1 + 2x_2 + 4x_3 \end{cases}$$, assim

$$T^{-1}(\mathbf{y}) = (-5y_1 + 2y_2 + 4y_3, 2y_1 - y_2 - y_3, y_1 - y_3)^T$$

Simplesmente poderíamos considerar inversão da matriz da transformação

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 1 & 3 \\ 1 & 2 & 1 \end{pmatrix} \to A^{-1} = \begin{pmatrix} -5 & 2 & 4 \\ 2 & -1 & -1 \\ 1 & 0 & -1 \end{pmatrix}$$, e escrever

$$\begin{aligned} T^{-1}(T(\mathbf{x})) = T^{-1}(A\mathbf{x}) &= T^{-1}(\mathbf{y}) \\ &= A^{-1}\mathbf{y} \\ &= \begin{pmatrix} -5 & 2 & 4 \\ 2 & -1 & -1 \\ 1 & 0 & -1 \end{pmatrix}\mathbf{y} \\ &= \begin{pmatrix} -5y_1 + 2y_2 + 4y_3 \\ 2y_1 - y_2 - y_3 \\ y_1 - y_3 \end{pmatrix} \end{aligned}$$

Obtendo mais uma vez a expressão da transformação inversa.

Podemos também escrever a expressão da inversa obtida em forma padrão, trocando y por x, para obter

$$T^{-1}(\mathbf{x}) = (-5x_1 + 2x_2 + 4x_3, 2x_1 - x_2 - x_3, x_1 - x_3)^T.$$

Observações:

a) Em todos os procedimentos utilizados, o custo computacional foi praticamente o mesmo. Demandou a resolução de um sistema linear de equações ou o cálculo da inversa de uma matriz.

b) Um resultado que ainda não foi provado formalmente e que nos permitiria mostrar a existência da transformação inversa é o seguinte: "Se a matriz de representação for inversível, então existe a inversa da transformação".

Por este motivo (e muitos outros que aparecerão ainda), a representação matricial de uma transformação linear se torna um dos problemas básicos da Álgebra Linear.

4.5 Representação Matricial de Transformações Lineares

Na seção anterior, mostramos que dada uma matriz $A_{m\times n}$ é possível definir uma transformação linear $T:\mathbb{R}^n \to \mathbb{R}^m$ dada por $T(\mathbf{x}) = A\mathbf{x}$ para cada $\mathbf{x} \in \mathbb{R}^n$.

Agora, mostraremos que para cada transformação linear $T:\mathbb{R}^n \to \mathbb{R}^m$ existe uma matriz $A_{m\times n}$ tal que $T(\mathbf{x}) = A\mathbf{x}$. Veremos também que qualquer operador linear definido entre dois espaços vetoriais de dimensão finita pode ser representado por uma matriz.

Teorema. *Seja $T:\mathbb{R}^n \to \mathbb{R}^m$, então existe uma matriz $A_{m\times n}$ tal que $T(\mathrm{x}) = A\mathrm{x}$ para cada $\mathrm{x} \in \mathbb{R}^n$. De fato, o j-ésimo vetor coluna da matriz A é dado por*

$$\mathbf{a}_j = T(\mathbf{e}_j), \quad j = 1, 2, \ldots, n$$

Demonstração. Seja a matriz $A_{m\times n}$ (explicitada através de suas colunas):

$$A = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_j, \ldots \mathbf{a}_n), \text{ onde } \mathbf{a}_j = (\mathbf{a}_{1j}, \mathbf{a}_{2j}, \ldots, \mathbf{a}_{mj})^T.$$

A matriz tem n colunas e cada coluna é um vetor com m componentes.

Para $j = 1, 2, \ldots, n$ definimos cada coluna como

$$\mathbf{a}_j = (\mathbf{a}_{1j}, \mathbf{a}_{2j}, \ldots, \mathbf{a}_{mj})^T = T(\mathbf{e}_j).$$

E se $\mathbf{x} \in \mathbb{R}^n$ for um vetor arbitrário escrito na base canônica

$$\boldsymbol{x} = (x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_n\boldsymbol{e}_n).$$

Fazendo $T(\mathbf{x})$ temos:

$$\begin{aligned} T(\mathbf{x}) &= x_1T(\mathbf{e}_1)+x_2T(\mathbf{e}_2)+\ldots+x_nT(\mathbf{e}_n) \\ &= x_1\mathbf{a}_1+x_2\mathbf{a}_2+\ldots+x_n\mathbf{a}_n \\ &= (\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= A\mathbf{x}. \end{aligned}$$

■

Mostramos que cada transformação linear $T:\mathbb{R}^n \to \mathbb{R}^m$ pode ser representada em termos de uma matriz. O teorema mostra como construir a matriz que representa um determinado operador linear.

Para obter a primeira coluna de A, aplique o operador linear T ao primeiro vetor da base de $\mathbb{R}^n$.

Repita o procedimento até obter todas as colunas de A.

Como usamos os elementos da base canônica de $\mathbb{R}^n$, $\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n$, dizemos que A é a representação canônica de T ou a matriz de T em relação às bases canônicas. Veremos, a seguir, como representar um operador linear em relação a outras bases.

Exemplo 28. Encontre a transformação $T:\mathbb{R}^3 \to \mathbb{R}^2$ dada por $T(\mathbf{x}) = (x_1+x_2, x_2+x_3)^T$ para cada $\mathbf{x} \in \mathbb{R}^n$.

Solução. Queremos encontrar uma matriz A tal que $T(\mathrm{x}) = A\mathrm{x}$. Para fazer isso, construiremos a matriz obtendo suas colunas consecutivamente calculando $\mathbf{a}_1 = T(\mathbf{e}_1)$, $\mathbf{a}_2 = T(\mathbf{e}_2)$ e $\mathbf{a}_3 = T(\mathbf{e}_3)$.

Como segue

$$\mathbf{a}_1 = T(\mathbf{e}_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\mathbf{a}_2 = T(\mathbf{e}_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\mathbf{a}_3 = T(\mathbf{e}_3) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

E então escolhemos esses vetores como as colunas de A,

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Podemos conferir o resultado calculando $A\mathbf{x}$

$$A\mathbf{x} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \end{pmatrix}.$$

Acabamos de ver como matrizes representam transformações lineares de $\mathbb{R}^n$ em $\mathbb{R}^m$, poderíamos perguntar agora se é possível encontrar uma representação análoga para operadores lineares de V em W, onde V e W são espaços de dimensão n e m, respectivamente. Veja o seguinte exemplo.

Veremos que em geral a matriz de representação de uma transformação linear depende das bases escolhidas para o domínio e a imagem. Faremos uma generalização do teorema anterior.

Para fazer isso, consideremos $E = [\mathbf{v}_1, \mathbf{v}_2 ..., \mathbf{v}_n]$ e $F = [\mathbf{w}_1, \mathbf{w}_2 ..., \mathbf{w}_m]$ bases ordenadas para V e W, respectivamente e uma transformação linear $T : V \to W$.

Se v é um vetor arbitrário em V podemos expressá-lo em termos da base E como:

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n.$$

Vamos mostrar que existe uma matriz $A_{m\times n}$ que representa o operador T no seguinte sentido:

$$A\mathbf{x} = \mathbf{y} \quad \text{se e somente se } T(\mathbf{v}) = y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \ldots + y_m\mathbf{w}_m.$$

A matriz A deve caracterizar o efeito do operador T.

Prova. Pela linea ridade do nosso operador T, temos:

$$T(\mathbf{v}) = x_1 T(\mathbf{v}_1) + x_2 T(\mathbf{v}_2) + \ldots + x_n T(\mathbf{v}_n).$$

Para cada $k, (1 \le k \le n)$, $T(\mathbf{v}_\mathbf{k})$ está em W e então pode ser representado em forma única em termos dos vetores da base F de W como: $T(\mathbf{v}_k) = a_{1k}\mathbf{w}_1 + a_{2k}\mathbf{w}_2 + \ldots + a_{mk}\mathbf{w}_m$.

A notação utilizada aqui para os coeficientes da combinação linear aik nos ajudará a enxergar a matriz que estamos procurando.

Assim,

$$\begin{aligned} T(\mathbf{v}) &= \\ &= x_1(a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \ldots + a_{m1}\mathbf{w}_m) \\ &+ x_2(a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \ldots + a_{m2}\mathbf{w}_m) \\ &+ \cdots \\ &+ x_n(a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \ldots + a_{mn}\mathbf{w}_m). \end{aligned}$$

Remanejando os termos apropriadamente, temos

$$\begin{aligned} T(\mathbf{v}) &= \\ &= (a_{11}\mathbf{x}_1 + a_{12}\mathbf{x}_2 + \ldots + a_{1n}\mathbf{x}_n)\mathbf{w}_1 \\ &+ (a_{21}\mathbf{x}_1 + a_{22}\mathbf{x}_2 + \ldots + a_{2n}\mathbf{x}_n)\mathbf{w}_2 \\ &+ \cdots \\ &+ (a_{m1}\mathbf{x}_1 + a_{m2}\mathbf{x}_2 + \ldots + a_{mn}\mathbf{x}_n)\mathbf{w}_m. \end{aligned}$$

Podemos enxergar os coeficientes de $\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_m$ na última expressão, como sendo exatamente o produto das linhas de A pelo vetor $\mathbf{x} = (x_1, x_1, \ldots x_n)^T$

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A\mathrm{x}$$

a coluna de A, (denotada por a_k, $1 \leq k \leq n$) é exatamente:

$$\mathbf{a}_k = (T(\mathbf{v}_k))_F = \begin{pmatrix} x_{1k} \\ x_{2k} \\ \vdots \\ a_{mk} \end{pmatrix}.$$

Então, a matriz A, cujas colunas são as coordenadas dos vetores $T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)$, é a matriz desejada. A matriz é única para as bases escolhidas desde que as coordenadas dos vetores sejam únicas nessas bases.

Por outro lado também temos que o vetor de coordenadas de $T(\mathbf{v})$ em relação a $[\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m]$ são da forma $y_i = \sum_{j=1}^{n} a_{ij}x_j$. Logo $\mathbf{y} = A\mathbf{x}$.

Acabamos de mostrar o seguinte teorema:

Teorema. *Seja uma transformação linear* $T: V \to W$ *e duas bases ordenadas* $E = [\mathbf{v}_1, \mathbf{v}_2 ..., \mathbf{v}_n]$ *e* $F = [\mathbf{w}_1, \mathbf{w}_2 ..., \mathbf{w}_m]$ *para* V *e* W *respectivamente, então existe uma matriz* $A_{m \times n}$ *tal que* $[T(\mathbf{v})]_F = A[\mathbf{v}]_E$. *Seja* A *a matriz de* $T(\mathrm{v})$ *em relação as bases ordenadas* E *e* F. *De fato, o j-ésimo vetor coluna da matriz* A *é dado por* $a_j = [T(\mathbf{v}_j)]_F$.

Se denotarmos por

- $\mathbf{x} = [\mathbf{v}]_E \in \mathbb{R}^n$: vetor de coordenadas de v em relação a E.
- $\mathbf{y} = [\mathbf{w}]_F \in \mathbb{R}^m$: vetor de coordenadas de w em relação a F.

Temos que

$$T: V \to W \text{ se e somente se } A\mathbf{x} = \mathbf{y}.$$

Mais uma vez as colunas de A são

$$A = (\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n) = ([T(\mathbf{v}_1)]_F, [T(\mathbf{v}_2)]_F, ..., [T(\mathbf{v}_n)]_F).$$

Para encontrar a representação matricial de uma TL $T: \mathbb{R}^n \to \mathbb{R}^m$ em relação às bases ordenadas $E = [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n]$ e $F = [\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_m]$, precisamos representar cada vetor $T(\mathrm{u}_j)$ como uma combinação linear de $\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n$. O teorema a seguir mostra que encontrar essa representação é equivalente a resolver o sistema linear $B\mathbf{x} = T(\mathbf{u}_j)$.

Teorema. *Sejam* $E = [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n]$ *e* $F = [\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_m]$ *bases ordenadas para* $\mathbb{R}^n$ *e* $\mathbb{R}^m$, *respectivamente. Se* $T: \mathbb{R}^n \to \mathbb{R}^m$ *é uma transformação linear e* A *é a matriz de* T *em relação a* E *e* F, *então* $\mathbf{a}_j = B^{-1}T(\mathbf{u}_j)$, *onde* $B = (\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_m)$.

Demonstração. Se A é matriz de representação de T em relação a E e F, então para $j = 1, 2, \ldots, n$, temos

$$\begin{aligned} T(\mathbf{u}_j) &= a_{1j}\mathbf{b}_1 + a_{2j}\mathbf{b}_2 + \ldots + a_{mj}\mathbf{b}_m \\ &= B\mathbf{a}_j. \end{aligned}$$

A matriz B é inversível, já que suas colunas formam uma base para $\mathbb{R}^m$, portanto $\mathbf{a}_j = B^{-1}T(\mathbf{u}_j)$.

■

Uma conseqüência desse teorema é que podemos determinar a matriz associada a um operador calculando a forma escalonada reduzida por linhas de uma matriz aumentada. O próximo corolário mostra como fazer isso.

Corolário. Se A é a matriz de representação de um operador linear $T:\mathbb{R}^n \to \mathbb{R}^m$ em relação as bases $E=[\mathbf{u}_1,\mathbf{u}_2,...,\mathbf{u}_n]$ e $F=[\mathbf{b}_1,\mathbf{b}_2,...,\mathbf{b}_m]$, então a forma escalonada reduzida por linhas da matriz

$$(\mathbf{b}_1,\mathbf{b}_2,...,\mathbf{b}_m)\,|\,T(\mathbf{v}_1),T(\mathbf{v}_2),...,T(\mathbf{v}_n)) \text{ é } (I\,|\,A)$$

Demonstração. Seja $B=(\mathbf{b}_1,\mathbf{b}_2,...,\mathbf{b}_m)$. A matriz

$$(B\,|\,T(\mathbf{v}_1),T(\mathbf{v}_2),...,T(\mathbf{v}_n))$$

é equivalente por linhas a

$$\begin{aligned}&B^{-1}(B\,|\,T(\mathbf{v}_1),T(\mathbf{v}_2),...,T(\mathbf{v}_n))=(I\,|\,B^{-1}T(\mathbf{v}_1),B^{-1}T(\mathbf{v}_2),...,B^{-1}T(\mathbf{v}_n))\\&=I\,|\,\mathbf{a}_1,\mathbf{a}_2,...,\mathbf{a}_n)\\&=(I\,|\,A).\end{aligned}$$

■

Exemplo 29. Seja a TL $T:\mathbb{R}^2 \to \mathbb{R}^3$ definida por

$$T(\mathbf{x})=(x_2,x_1+x_2,x_1-x_2)^T.$$

Encontre a matriz de T em relação as bases ordenadas

$$E=[\mathbf{u}_1,\mathbf{u}_2]=[(1,2)^T,(3,1)^T \text{ e}$$

$$F=[\mathbf{b}_1,\mathbf{b}_2\mathbf{b}_3]=[(1,0,0)^T,(1,1,0)^T,(1,1,1)^T].$$

Solução. O procedimento de solução consiste em formar a matriz aumentada e depois levá-la a forma escalonada reduzida por linhas. Para isso, precisamos calcular previamente $T(\mathbf{v}_1),T(\mathbf{v}_2)$; assim temos que:

$$(\mathbf{b}_1,\mathbf{b}_2\mathbf{b}_3\,|\,T(\mathbf{v}_1),T(\mathbf{v}_2))=\begin{pmatrix}1&1&1&|&2&1\\0&1&1&|&3&4\\0&0&1&|&-1&2\end{pmatrix}\to\begin{pmatrix}1&0&0&|&-1&-3\\0&1&0&|&4&2\\0&0&1&|&-1&2\end{pmatrix}$$

Logo, a matriz procurada é $A = \begin{pmatrix} -1 & -3 \\ 4 & 2 \\ -1 & 2 \end{pmatrix}$

O leitor pode verificar que

$$T(u_1) = -b_1 + 4b_2 - b_3$$
$$T(u_2) = -3b_1 + 2b_2 + 2b_3 \, .$$

Exercícios

1) Para cada uma das transformações lineares T no Exercício 1 da Seção 1, encontre a matriz A que representa T.

2) Para cada uma das transformações lineares L de $\mathbb{R}^3$ em $\mathbb{R}^2$ a seguir, encontre uma matriz A, tal que $T(\mathbf{x}) = A\mathbf{x}$ para todo $\mathbf{x}$ em $\mathbb{R}^3$.

   a) $T((x_1, x_2, x_3)^T) = (x_1 + x_2, 0)^T$

   b) $T((x_1, x_2, x_3)^T) = (x_1, x_2)^T$

   c) $T((x_1, x_2, x_3)^T) = (x_2 - x_1, x_3 - x_2)^T$

3) Para cada uma das transformações lineares T de $\mathbb{R}^3$em $\mathbb{R}^3$ a seguir, encontre uma matriz A, tal que $T(\mathbf{x}) = A\mathbf{x}$ para todo $\mathbf{x}$ em $\mathbb{R}^3$.

   a) $T((x_1, x_2, x_3)^T) = (x_3, x_2, x_1)^T$

   b) $T((x_1, x_2, x_3)^T) = (x_1, x_1 + x_2, x_1 + x_2 + x_3)^T$

   c) $T((x_1, x_2, x_3)^T) = (2x_3, x_2 + 3x_1, 2x_1 - x_3)^T$

4) Seja L a transformação linear de $\mathbb{R}^3$ em $\mathbb{R}^3$ definida por $T(x) = (2x_1 - x_2 - x_3, 2x_2 - x_1 - x_3, 2x_3 - x_1 - x_2)$. Determine a matriz A de L em relação a base canônica e use-a para encontrar $T(\mathbf{x})$ para cada um dos vetores $\mathbf{x}$ a seguir.

   a) $\mathbf{x} = (1,1,1)^T$

   b) $\mathbf{x} = (2,1,1)^T$

   c) $\mathbf{x} = (-5,3,2)^T$

5) Encontre a representação matricial canônica para cada um dos operadores lineares T em $\mathbb{R}^2$ descritos a seguir.

a) T roda cada vetor $\mathbf{x}$ de 45° no sentido anti-horário.

b) T reflete cada vetor $\mathbf{x}$ em relação ao eixo dos x_1 e depois roda o vetor refletido de 90° no sentido trigonométrico.

c) T dobra o comprimento do eixo, depois roda o vetor obtido de 30° no sentido trigonométrico.

d) T reflete cada vetor $\mathbf{x}$ em relação a reta $x_1 = x_2$ e depois projeta o vetor refletido sobre o eixo dos x_1.

6) Sejam $\mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{b}_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{b}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ e seja T a transformação linear de $\mathbb{R}^2$ em $\mathbb{R}^3$ definida por $T(\mathbf{x}) = x_1 b_1 + x_2 b_2 + (x_1 + b_2) b_3$. Encontre a matriz A de T em relação às bases $[e_1, e_2]$ e $[b_1, b_2, b_3]$.

7) Sejam $y_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $y_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $y_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ e seja I o operador identidade em $\mathbb{R}^3$.

a) Encontre as coordenadas de $I(\mathbf{e}_1), I(\mathbf{e}_2), I(\mathbf{e}_3)$ em relação a $[y_1, y_2, y_3]$.

b) Encontre uma matriz A tal que $A\mathbf{x}$ é o vetor de coordenadas de $\mathbf{x}$ em relação a $[y_1, y_2, y_3]$.

8) Sejam y_1, y_2, y_3, dados no exercício anterior e seja T a transformação linear de $\mathbb{R}^3$ em $\mathbb{R}^3$ definida por:

$$T(c_1 y_1 + c_2 y_2 + c_3 y_3) = (c_1 + c_2 + c_3) y_1 + (2c_1 + c_3) y_2 - (2c_2 + c_3) y_3$$

$$T(c_1 y_1 + c_2 y_2 + c_3 y_3) = (c_1 + c_2 + c_3) y_1 + (2c_1 + c_3) y_2 - (2c_2 + c_3) y_3$$

a) Encontre a matriz de T em relação à base ordenada $[y_1, y_2, y_3]$.

b) Escreva cada um dos vetores $\mathbf{x}$ a seguir como uma combinação linear de y_1, y_2, y_3 e use a matriz encontrada em a) para determinar $T(\mathbf{x})$.

i) $\mathbf{x} = (7, 5, 2)^T$ ii) $\mathbf{x} = (3, 2, 1)^T$ iii) $\mathbf{x} = (1, 2, 3)^T$

9) Sejam $E = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$ e $F = [\mathbf{b}_1, \mathbf{b}_2]$, onde $\mathbf{u}_1 = (1, 0, -1)^T$, $\mathbf{u}_2 = (1, 2, 1)^T$, $\mathbf{u}_3 = (-1, 1, 1)^T$ e $\mathbf{b}_1 = (1, -1)^T$, $\mathbf{b}_2 = (2, -1)^T$. Para cada uma das transformações lineares T de $\mathbb{R}^3$ em $\mathbb{R}^2$ a seguir, encontre a matriz de T em relação às bases ordenadas E e F.

a) $T(\mathbf{x}) = (x_3, x_1)^T$

b) $T(\mathbf{x}) = (x_1 + x_2, x_1 - x_3)^T$

c) $T(\mathbf{x}) = (2x_2, -x_1)^T$

4.6 Matrizes e Transformações Lineares, Equivalências e Propriedades

Seja T uma transformação linear de V em W ($\dim(V) = \dim(W) = n$) e A a matriz de representação de T em relação as bases de V e de W. As afirmações a seguir são equivalentes:

i) A é inversível, isto é, existe A^{-1}

ii) $A\mathrm{x} = \mathrm{b}$ tem solução única

iii) $A\mathrm{x} = 0$ só admite a solução trivial

iv) A forma escalonada reduzida por linhas de A é a matriz identidade

v) A é um produto de matrizes elementares

vi) $\text{posto}(A) = n$

vii) $\text{nulidade}(A) = 0$

viii) A tem n vetores coluna linearmente independentes

ix) O espaço coluna de A gera $\mathbb{R}^n$

x) O espaço linha de A gera $\mathbb{R}^n$

xi) O espaço nulo de A é o vetor nulo

xii) $\det(A) \neq 0$

xiii) T é inversível e a matriz associada com a transformação inversa T^{-1} é A^{-1}

xiv) T é injetora e sobrejetora

xv) T é um isomorfismo

xvi) $\ker(T) = \{0\}$, $\mathrm{Im}(T) = W$

O conceito de transformações lineares, um tópico de extrema importância da Matemática, tem sido definido e analisado à luz de seu núcleo (kernel), imagem e também através da matriz associada à transformação.

Achar a matriz de representação de uma transformação linear pode ser considerado como um dos principais problemas básicos da Álgebra Linear.

Os conceitos aprendidos nesta primeira disciplina são ferramentas matemáticas básicas da Álgebra Linear. Com uma boa familiaridade em álgebra matricial, transformações lineares e espaços vetoriais, estaremos prontos para compreender e enfrentar outros importantes problemas da Álgebra Linear como: o conceito de ortogonalidade, o estudo de auto-sistemas (autovalores e autovetores de uma matriz) e o problema de diagonalização; tópicos correspondentes à disciplina de *Álgebra Linear II*.

Bibliografia Comentada

LAY, David C. *Álgebra linear e suas aplicações*. 2. ed. [S.l]:LTC, [200-?].

O texto fornece uma introdução elementar e moderna da álgebra linear e algumas de suas aplicações interessantes, acessível a alunos com a maturidade que dois semestres completos de matemática em nível de terceiro grau, em disciplinas de cálculo em geral, lhes conferem. O objetivo é ajudar os alunos a dominar os conceitos e habilidades básicos que usarão mais tarde em suas carreiras. Os tópicos escolhidos seguem as recomendações do Linear Algebra Curriculum Study Group, que, por sua vez, baseiam-se em uma cuidadosa pesquisa sobre as necessidades reais dos alunos e em um consenso entre os profissionais dos muitos campos que usam a Álgebra Linear.

POOLE, David. *Álgebra linear*. São Paulo: Pioneira Thomson Learning, 2004.

Este livro foi estruturado de forma bastante flexível, com a preocupação central de que a álgebra linear constitua um assunto estimulante o suficiente e de fácil aprendizado, tornando-a mais acessível ao estudante. Escrito de forma clara, direta e objetiva, aborda temas como Vetores, Matrizes, Autovalores e Autovetores, Ortogonalidade, Espaços Vetoriais, e Distância e Aproximação. A apresentação de conceitos-chave com antecedência, a ênfase em vetores e geometria e os inúmeros exercícios e exemplos que reforçam o fato de a Álgebra Linear ser uma ferramenta valiosa para a modelagem de problemas da vida real consistem no principal diferencial deste livro. A apresentação de pequenos esboços biográficos de muitos dos matemáticos que contribuíram para o desenvolvimento da Álgebra Linear é outro diferencial, valorizando a história da matemática.

STEVE, Leon J. *Álgebra linear com aplicações*. 4. ed. [S.l.]: LTC, [200-?].

Este livro é apropriado para alunos que tenham conceitos básicos de matrizes e tenham passado por um curso de Geometria Analítica. O estudante deve estar também familiarizado com as noções básicas de Cálculo Diferencial e Integral. Esta nova edição, ao mesmo tempo que mantém a essência das edições anteriores, incorpora uma série de melhorias substanciais: - Conjunto de Exercícios Computacionais em cada Capítulo; - Mais Motivação Geométrica; - Nova Aplicação Envolvendo Teoria dos Grafos e Redes; - Motivação Adicional para a Definição de Determinantes; - A seção sobre Mudança de Base foi transferida para o Cap. 3; - Revisões Importantes na seção sobre Espaços Munidos de Produto Interno; - A seção sobre Normas Matriciais foi transferida para o Cap. 7; - Nova Aplicação: Aproximação de Funções por Polinômios Trigonométricos; - Revisões no Cap. 6.